تأثير ما وراء البيانات

على كفاءة استرجاع مقالات الدوريات الالكترونية العربية



تأليف

د. منيرة محمد مظهر

مدرس المكتبات والمعلومات

كلية الآداب - جامعة القاهرة

تقديم

أ. د. حشمت قاسم

أستاذ علم المعلومات المتفرغ

كلية الآداب - جامعة القاهرة

دار الفجر للنشر والتوزيع

تأثير ما وراء البيانات

على كفاءة استرجاع

مقالات الدوريات الإلكترونية العربية

تأثير ما وراء البيانات
على كفاءة استرجاع مقالات الدوريات الإلكترونية العربية

الطبعة الأولى 2018

تأليـف: د. منيرة محمد مظهر لطفي

رقم الإيداع : 15056

ISBN : 978-977-358-377-4

الناشر

دار الفجر للنشر والتوزيع

4 شارع هاشم الأشقر - النزهة الجديدة

القاهرة - مصر

تليفون: 26242520 - 26246252(00202)

فاكس: 26246265 (00202)

Email: info@daralfajr.com

تأثير ما وراء البيانات على كفاءة استرجاع مقالات الدوريات الإلكترونية العربية

تأليف

د. منيرة محمد مظهر لطفي

مدرس المكتبات والمعلومات

كلية الآداب - جامعة القاهرة

تقديم

الأستاذ الدكتور حشمت قاسم

أستاذ علم المعلومات المتفرغ

كلية الآداب - جامعة القاهرة

دار الفجر للنشر والتوزيع

2018

بسم الـلـه الرحمن الرحيم

(وَهُوَ الـلـهُ لَا إِلَهَ إِلَّا هُوَ لَهُ الْحَمْدُ فِي الْأُولَى وَالْآخِرَةِ وَلَهُ الْحُكْمُ وَإِلَيْهِ تُرْجَعُونَ {70/28})

صدق الـلـه العظيم

(سورة القصص، الآية 70)

الإهداء

إلى اللذين لولاهما

- بعد الـلـه عز وجل -

ما كان هذا العمل وما كنت أنا شيئا

إلى والديَ

جزاهما الـلـه عني خير الجزاء

والحمد لله رب العالمين

شكر وتقدير

الحمد لله رب العالمين حمدا كثيرا طيبا مباركا فيه كما ينبغي لجلال وجهه وعظيم سلطانه وفضله، والصلاة والسلام على خير خلق اللـه سيدنا محمد وعلى آله وصحبه أجمعين.

في هذا المقام لا يسعني إلا أن أتقدم بخالص الشكر والتقدير والعرفان إلى أستاذي العالم الجليل الأستاذ الدكتور/ حشمت محمد علي قاسم أستاذ علم المعلومات المتفرغ بكلية الآداب جامعة القاهرة الذي شرفني بتقديم هذا العمل الذي جاء ثمرة لأطروحة الدكتوراه التي تفضل سيادته بالإشراف عليها، ولا أملك إلا أن أسأل اللـه عز وجل أن يجازيه عني خير الجزاء، وأن يزيده من علمه وواسع فضله.

كما أتقدم بجزيل الشكر وعظيم التقدير والاحترام لأستاذي العالم الفاضل الدكتور/ مصطفى أمين حسام الدين الأستاذ المساعد المتفرغ لعلم المعلومات بكلية الآداب جامعة القاهرة الذي قدم لي من العون والإرشاد والدعم ما كان له عظيم الأثر في نفسي وفي إنجاز هذا العمل، فجزاه اللـه عني خير الجزاء.

وحسبي أنني بشر أصيب وأخطئ، فإن أصبت فبتوفيق من اللـه عز وجل، وإن أخطأت فمني.

(وَمَا تَوْفِيقِي إِلاَّ بِاللـهِ عَلَيْهِ تَوَكَّلْتُ وَإِلَيْهِ أُنِيبُ {88/11}) [هود: 88]

المحتويات

قائمة الجداول

قائمة الأشكال

تقديم

تنظيم المعلومات في بيئة العنكبوتية العالمية واحد من أهم موضوعات علم المكتبات والمعلومات المعاصر. يعنى هذا الموضوع بالبحث في المقاربات النظرية والعملية والأساليب والتقنيات التي تسهم في وصف وتكشيف مصادر المعلومات المتاحة على العنكبوتية العالمية، والدراسة المنهجية للنظم والأدوات والقنوات المتعلقة بتيسير سبل البحث عنها واكتشافها واسترجاعها، وكذلك للمعايير والتقنينات التي ينبغي الالتزام بها لتحقيق كل ذلك بجودة وكفاءة وفعالية.

ونظرا للدور الرئيس الذي تؤديه ما وراء البيانات "الميتاداتا" في تنظيم المعلومات في بيئة العنكبوتية العالمية، فقد كانت وما زالت محورا للعديد من البحوث والدراسات العربية التي تتناول قضايا تتعلق بخصائصها ووظائفها ولغات التكويد والمعايير المستخدمة في إنشائها، الخ. ومن بين أهم هذه القضايا، القضية التي يتناولها هذا الكتاب وهي قياس تأثير ما وراء البيانات على كفاءة استرجاع مصادر المعلومات العربية من خلال محركات البحث على العنكبوتية العالمية.

وتطبيقا لهذه القضية، اختارت الباحثة مقالات الدوريات العربية الالكترونية في مجال المكتبات والمعلومات، باعتبار أن المقالات أحد أهم مصادر المعلومات في إطار الاتصال العلمي بين المشتغلين في هذا المجال، كما اختارت المنهج التجريبي للتحقق من الفروض التي وضعتها لضبط مسار الدراسة والوصول إلى نتائج موضوعية ذات دلالة كمية محسوسة قابلة للمقارنة.

ويكشف الجهد العلمي الأصيل، الذي بذلته الباحثة في تناول هذا الموضوع الفني بالغ الأهمية والحداثة، عن معالجة محكمة تتميز بالدقة والوضوح والإيجاز، واستخدام واع للإحصاء الوصفي والاستدلالي، فضلا عن أمانة في التوثيق.

إن الزاد المعرفي، المبني على أسس منهجية الذي يقدمه هذا الكتاب، والذي قدم أصلا كأطروحة لنيل درجة الدكتوراه في تقنيات المعلومات من كلية الآداب-جامعة القاهرة، يسهم في تنمية الوعي بمدى تأثير ما وراء البيانات المعيارية على تحسين كفاءة استرجاع مصادر المعلومات العربية، والمساعدة في تطوير استخدامها واستثمارها في قواعد البيانات والمكتبات والمستودعات الرقمية العربية. وهو إضافة ثرية للمكتبة العربية في الموضوع نرجو أن يفيد منه المعنيون بالعنكبوتية العالمية ومحركات البحث بصفة عامة، والمعنيون بمصادر المعلومات العربية على العنكبوتية العالمية إنشاء وتصميما وبثا، تنظيما وتكشيفا وإتاحة، بحثا واسترجاعا وإفادة، وكذلك المعنيون بعمليات الرقمنة ونظم استرجاع المعلومات في الوطن العربي بصفة خاصة.

والله من وراء القصد،

الأستاذ الدكتور / حشمت قاسم

مقدمة

ظهرت ما وراء البيانات في بيئة العنكبوتية كوسيلة لتسهيل العثور على المصادر واسترجاعها، وتعتمد على وصف المصادر الإلكترونية من خلال تيجان Tags تسجل في جزء الرأس head من شفرة المصدر Source Code لصفحة العنكبوتية Web Page، وهذه التيجان لا يراها المستفيد النهائي، ومن هنا جاءت التسمية بما وراء البيانات.

ومنذ ظهور ما وراء البيانات وهي تشكل مجالا خصبا للدراسة والبحث من قبل تخصصات مختلفة منها المكتبات والمعلومات، والحاسبات الآلية، وتصميم المواقع، وتحسين أداء محركات البحث Search Engines Optimization (SEO). كما تعد ما وراء البيانات أحد الموضوعات التي تحظى بمزيد من الاهتمام في الوقت الراهن باعتبارها أحد العناصر الهامة في التوجه نحو العنكبوتية الدلالية Semantic Web. ولذلك فإن هذه الدراسة تركز على ما وراء البيانات في مقالات الدوريات الإلكترونية العربية في مجال المكتبات والمعلومات، ومدى تأثير توافرها على كفاءة استرجاعها من محركات البحث.

وتنبع مشكلة الدراسة وأهميتها من عدد من الملاحظات تتمثل في:

- اختلاف وتضارب الدراسات حول مكانة وأهمية ما وراء البيانات، و تأثيرها على محركات البحث، فعلى الرغم من وجود عدد من الدراسات التي تؤكد على أهمية ما وراء البيانات ودورها في استرجاع المصادر الإلكترونية للمعلومات من محركات البحث والتوصية بضرورة استخدام بؤرة دبلن[1] توجد في المقابل دراسات أخرى

(1) نذكر من هذه الدراسات على سبيل المثال:

عبد الهادي، زين. وصائف البيانات (Metadata) في مواقع المكتبات العربية في كل من مصر والسعودية [على الخط المباشر]: دراسة تطبيقية. Cybrarians Journal، ع 4 (مارس 2005). متاح على: http://www.journal.cybrarians.info/index.php?option=com_content&view=article&id=503:metadata&catid=241:2011-08-16-08-16-55 . [تاريخ الزيارة 2015/4/2].

Alimohammadi, D. Measurement of the Presence of keywords and description meta-tags on a selected number of Iranian Websites . Online Information Review, Vol. 28, No.3 (2004). pp. 220-223.

تشير إلى ضعف تأثير ما وراء البيانات على محركات البحث[1]، والذي جاء كنتيجة لبعض الممارسات الخاطئة من قبل مصممي المواقع تمثلت في إدراج عدد كبير من الكلمات المفتاحية التي لا علاقة لها بمحتويات مواقعهم في تيجان ما وراء البيانات بهدف رفع ترتيب هذه المواقع في قوائم نتائج محركات البحث.

- اختلاف توجه محركات البحث فيما بينها نحو ما وراء البيانات؛ فبعض محركات البحث مثل جوجل Google أعلن عن تجاهله لعناصر ما وراء البيانات، وعدم استخدامه لها، بينما محركات بحث أخرى مثل ياهو Yahoo لا تزال تدعم عناصر ما وراء البيانات مثل عنصر الكلمات المفتاحية [2] يضاف إلى ذلك أنه على الرغم من إعلان محرك البحث جوجل Google عن تجاهله لما وراء البيانات وبخاصة في الترتيب الطبقي للمواقع في قائمة نتائج محرك البحث، إلا أنه قد أعلن عن استخدامه لبعض تيجان ما وراء البيانات في تقديم المقتطف الغني Rich Snippet للمواقع في قائمة نتائج البحث.
- التوجه نحو العنكبوتية الدلالية، والتي تعتمد بصفة أساسية على جعل البيانات على العنكبوتية مفهومة لكل من البشر والآلات عن طريق استخدام ما وراء البيانات.

---

(1) نذكر من هذه الدراسات على سبيل المثال:
Elser, J.K. & Paxton , J.. A Search Engine That Learns . In: Proceedings of the Fourth International Association of Science and Technology for Development (IASTED) International Conference on Computational Intelligence, Canada: Calgary, July 2005.pp. 316-320. Also: Konia, B.S. Search Engine Optimization with WebPosition GOLD 2. Plano, Texas: Wordware Publishing, 2002.

(2) Meta element [online]. In: Wikipedia: The Free Encyclopedia .Available at: https://en.wikipedia.org/wiki/Meta_element. [Accessed 11/9/2015].

- الأهمية الخاصة لما وراء البيانات بالنسبة لمقالات الدوريات الإلكترونية؛ والتي ترجع إلى طبيعة مقالات الدوريات وما تتميز به من خصائص. فمقالات الدوريات يعد كل منها مصدر معلومات مستقل بذاته له بياناته التي تختلف من مقالة لأخرى، وترد كل مقالة في صفحة مستقلة في موقع الدورية الإلكترونية التي نشرت بها، ويمكن أن يتم البحث عنها من خلال محركات البحث، أو في أحد قواعد البيانات – كما هو الحال في قواعد البيانات الضخمة التي تضم أعدادا كبيرة من الدوريات – أو في أحد المستودعات الرقمية، ووفقا لذلك فإن تسجيل ما وراء البيانات في كل مقالة من مقالات الدورية الإلكترونية بما يتناسب معها أمرا ضروريا وله أهمية خاصة فيما يتعلق باسترجاع هذا النوع الخاص من مصادر المعلومات.

وبناء على هذه الملاحظات، يتضح أن ما وراء البيانات وتأثيرها على محركات البحث يحتاج إلى مزيد من الدراسة، الأمر الذي يبرز أهمية الدراسة الحالية التي تحاول التعرف على ما وراء البيانات في مقالات الدوريات الإلكترونية العربية في مجال المكتبات والمعلومات، ومدى تأثير توافرها على استرجاع هذه المقالات من محركات البحث.

هذا وتهدف الدراسة إلى:

- التعرف على الوضع الحالي لما وراء البيانات في مقالات الدوريات الإلكترونية العربية في مجال المكتبات والمعلومات.
- تقديم نظرة استكشافية لما وراء البيانات في مقالات عدد من الدوريات الإلكترونية الأجنبية في مجال المكتبات والمعلومات.
- اختبار مدى كفاية عناصر ما وراء بيانات خطة بؤرة دبلن Dublin Core في وصف مقالات الدوريات الإلكترونية العربية في المجال بما يحقق وصفا مناسبا لهذه المقالات.
- اختبار مدى كفاية مفردات سكيما Schema.org في وصف مقالات الدوريات الإلكترونية العربية في المجال باستخدام إطار وصف المصدر في الخصائص RDFa بما يحقق وصفا مناسبا لهذه المقالات.

- اختبار تأثير توافر ما وراء البيانات على كفاءة استرجاع مقالات الدوريات الإلكترونية العربية في مجال المكتبات والمعلومات من محركات البحث.
- التحقق من إمكانية استثمار ما وراء بيانات مقالات الدوريات الإلكترونية العربية في مجال المكتبات والمعلومات من خلال استخلاصها وإيداعها في أي نظام آلي يدعم شكل الاتصال مارك 21 (MARC 21).

وفي ضوء الأهداف التي تسعى الدراسة لتحقيقها، فإن التساؤلات التي تحاول الإجابة عليها هي:

1. ما الوضع الحالي لما وراء البيانات في مقالات الدوريات الإلكترونية العربية في مجال المكتبات والمعلومات ويشمل هذا التساؤل التساؤلات الفرعية التالية:
   - ما مدى توافر ما وراء البيانات في مقالات الدوريات الإلكترونية العربية في المجال؟
   - ما هي خطط ما وراء البيانات المستخدمة في مقالات الدوريات الإلكترونية العربية في المجال؟
   - ما أبرز تيجان (عناصر) ما وراء البيانات المستخدمة في مقالات الدوريات الإلكترونية العربية في مجال المكتبات والمعلومات؟
   - ما مدى جودة عناصر ما وراء البيانات الواردة في هذه المقالات؟
   - كيف تقوم الدوريات الإلكترونية الأجنبية في مجال المكتبات والمعلومات بإدراج عناصر ما وراء البيانات وما هي أكثر العناصر المستخدمة بها؟
2. ما مدى كفاية عناصر بؤرة دبلن في وصف مقالات الدوريات الإلكترونية العربية في مجال المكتبات والمعلومات؟
3. ما مدى كفاية مفردات سكيما Schema.org في وصف مقالات الدوريات الإلكترونية العربية المتخصصة في المجال باستخدام إطار وصف المصدر في الخصائص RDFa؟
4. هل يؤثر توافر ما وراء البيانات في مقالات الدوريات الإلكترونية العربية المتخصصة في المجال على كفاءة استرجاعها من محركات البحث؟

5. هل يمكن استثمار ما وراء البيانات في مقالات الـدوريات الإلكترونيـة العربيـة في المجـال واستخلاصها لإعداد تسجيلات تتوافق مع شكل الاتصال مارك 21 (MARC 21)؟

وقد اتخـذت الدراسـة عـددا مـن الحـدود لتعمـل في إطارهـا، ففـيما يتعلـق بالحـدود الموضـوعية فالدراسـة تتنـاول مـا وراء البيانـات في مقـالات الـدوريات الإلكترونيـة العربيـة المتخصصة في مجال المكتبات والمعلومات. أما الحدود النوعية فتقتصر الدراسة عـلى مقـالات الدوريات الإلكترونية العربية المتخصصة في مجـال المكتبـات والمعلومـات وتشـمل الـدوريات التاليـة: "Cybrarians Journal"، و "العربيـة 3000 [*]"، و"دراسـات المعلومـات"، و"المجلـة العربيـة للدراسـات المعلوماتيـة"، و"مجلـة المكتبـات والمعلومـات". وقـد أخـذت الدراسـة في الاعتبار صيغ عرض المقالات، حيث تم الاقتصار عـلى الـدوريات التـي تتـيح الـنص الكامـل أو الملخص لمقالاتها بلغة ترميز النص الفائق HTML. هذا فضلا عن استكشاف عنـاصر مـا وراء البيانـات في عينـة مـن مقـالات ثـلاث دوريـات إلكترونيـة أجنبيـة هـي كـل مـن: " D –Lib Magazine"، و"Ariadne"، و"First Monday". هـذا وتتنـاول الدراسـة مقـالات الـدوريات الإلكترونية العربية في مجال المكتبات والمعلومات بغض النظر عن لغة المقالة.

كما تقتصر الدراسة في حدودها الزمنيـة عـلى تنـاول المقـالات الصـادرة عـن الـدوريات الإلكترونية العربية المتخصصة في المجال والمتاحة عـلى موقـع كـل دوريـة حتـى ينـاير 2015، وهو تاريخ الانتهاء من تجميع المقالات للتحليل وتصميم التجربة.

ونظرا لأن هذه الدراسة تنقسم إلى قسمين أساسيين: القسم الأول يتعلق بتحليل ما وراء البيانات في مقالات الدوريات الإلكترونية في مجال المكتبات والمعلومات، والقسم الثاني يتعلق بتحديد مدى تأثير ما وراء البيانات على كفاءة استرجاع مقالات هذه الدوريات من محركـات البحث، فإن الدراسة تعتمد على كل من:

(*) توقف موقع الدورية لسبب غير معروف، ويمكن الاطلاع عليها عـلى أرشـيف الإنترنـت عـلى الـرابط https://web.archive.org/web/20160316204420/http://alarabiclub.org/index.php?p_id=212

المنهج الوصفي التحليلي: وتم الاستعانة بهذا المنهج في عرض ودراسة الواقع الفعلي لما وراء البيانات في مقالات الدوريات الإلكترونية في مجال المكتبات والمعلومات. وتم إعداد قائمة مراجعة لتحليل عناصر ما وراء البيانات الواردة في هذه المقالات يدويا، فضلا عن الاستعانة ببرنامج تحليل ما وراء البيانات ميتاشيكر[•] Metachecker.

المنهج التجريبي: تم الاعتماد على هذا المنهج في إجراء تجربة لاختبار مدى تأثير توافر ما وراء البيانات (المتغير التجريبي) على كفاءة الاسترجاع (المتغير التابع) لعينة من مقالات الدوريات العربية محل الدراسة، وذلك من محركات البحث: جوجل Google، وياهو Yahoo، وبينج Bing. كما تم الاعتماد على المنهج التجريبي في اختبار إمكانية استثمار عناصر ما وراء البيانات الواردة في المقالات من خلال استخلاصها وإعداد تسجيلات ببليوجرافية وفقا لشكل مارك 21 (MARC 21).

وللتعرف على واقع ما وراء البيانات في مقالات الدوريات الإلكترونية العربية في مجال المكتبات والمعلومات، تم تحديد عينة من مقالات الدوريات الخمسة المختارة. وتم اختيار هذه العينة بواقع ثلاث مقالات من كل عدد من الأعداد الصادرة عن كل دورية حتى يناير عام 2015 على أن تكون هذه المقالات أول ثلاث مقالات ترد في العدد بعد الافتتاحية إن وجدت. ووفقا لذلك فإن عينة الدراسة من المقالات بلغت (222) مقالة من إجمالي (653) مقالة وردت في (74) عددا صدرت عن هذه الدوريات. وهذه العينة بلغت درجة الثقة بها 95%، وبهامش خطأ 5.35%[1].

ولإجراء التجربة لقياس مدى تأثير ما وراء البيانات على محركات البحث، وبعد موافقة أربع دوريات فقط على إعادة نشر مقالاتها على الإنترنت ضمن دراسة تجريبية (Cybrarians

(•) تغيرت الإمكانات التي يوفرها هذا البرنامج عن الوقت الذي استخدم فيه لهذه الدراسة. ويمكن مراجعة الشكل القديم للبرنامج وإمكاناته على أرشيف الإنترنت على: https://web.archive.org/web/20160409144707/http://www.metachecker.net/

(1) تم حساب هذه النسبة باستخدام برنامج Sample Size Calculator المتاح على الموقع التالي: http://www.raosoft.com/samplesize.html

Journal، العربيـة 3000، دراسـات المعلومـات، المجلـة العربيـة للدراسـات المعلوماتيـة) تـم اختيار مقالة واحدة ممثلة لكل عدد من أعداد كل دورية مـن هـذه الـدوريات حتـى ينـاير 2015، ووفقا لذلك فقد بلغت عينة المقالات التي تم إجراء التجربة عليهـا (72 مقالـة) تمثـل 72 عددا هـي إجـمالي الأعـداد الصـادرة عـن الـدوريات محـل الدراسـة. وتـم تصـميم ثـلاث مجموعات تجريبية مجموعـة ضـابطة واحـدة تتكـون مـن مقـالات لا تشـتمل عـلى مـا وراء البيانات، ومجموعتين تجريبيتين تتكون الأولى من مقالات تشتمل على ما وراء البيانـات وفقـا لبؤرة دبلن، وتتكون الثانية من مقالات تشتمل على مـا وراء البيانـات وفقـا لمفـردات سـكيما باستخدام إطار وصف المصدر في الخصائص RDFa. وتم إتاحة هذه المجموعـات عـلى ثلاثـة مواقع في نفس الوقـت (16 ينـاير 2016)، وكانـت كالتـالي: موقـع www.lisaja.com لإتاحـة مقـالات المجموعـة الضـابطة - المجموعـة (أ)، وموقـع www.ajainlis.com لإتاحـة مقـالات المجموعة التجريبية الأولى – المجموعة (ب)، وموقع www.arabiclisja.com لإتاحة مقالات المجموعـة التجريبيـة الثانيـة - المجموعـة (ج)، ولضـمان تكشـيف محركـات البحـث لمواقـع التجربة تم إتاحة فترة زمنية مناسبة بلغت شهرين توقفت بعدها التجربة، وتلا ذلك البحـث عن عينة المقالات الخاضعة للتجربة بكل من العنوان، والكلمات المفتاحية.

إلا أنه تجدر الإشارة إلى أن هذه الدراسة واجهت صعوبات كان أبرزها الطبيعة المتغـيرة للإنترنت والتي تشكل أحد أبرز خصائص هذه الشبكة، وقد ظهر ذلك جليا في اختفاء وتوقف أحد مواقع الدوريات الإلكترونية محل الدراسة، وهي دورية العربية 3000، ويمكـن مراجعـة موقع الدورية عـلى أرشـيف الإنترنـت Internet Archive وهـو الموقـع الـذي يقـدم خدمـة أرشفة لصفحات ومواقع الإنترنت. كذلك تغير شكل وإمكانات برنامج تحليل ما وراء البيانـات ميتاشيكر Metachecker عن الوقت الذي استخدم فيه لهذه الدراسة.

وبمراجعة الدراسات السابقة حول ما وراء البيانات يتبـين أن الدراسـة الحاليـة تختلـف عن الدراسات العربية السابقة في أن أغلب الدراسات السابقة يتناول تحليل ما وراء البيانات من حيث مدى استخدام ما وراء البيانات في المواقع والتيجان والمعـايير المسـتخدمة والقليـل منها يدخل في إطار الدراسات التجريبية، كما أن هذه الدراسات ركـزت عـلى تحليـل عنـاصر

ما وراء البيانات على مستوى الموقع ككل، بينما تسعى الدراسة الحالية إلى تحليل ما وراء البيانات على مستوى الصفحات الداخلية للموقع الواحد وهي هنا مقالات الدوريات التي تشكل صفحات لموقع واحد هو موقع الدورية. كما أن الدراسة الحالية تتعرض للمفهوم الجديد لما وراء البيانات في بيئة العنكبوتية الدلالية Semantic Web في الجزء المتعلق بقياس تأثير ما وراء البيانات وفقا لمفردات سكيما Schema.org باستخدام إطار وصف المصدر في الخصائص RDFa على كفاءة الاسترجاع من محركات البحث.

وتنتظم الدراسة في ثلاثة فصول هي:

الفصل الأول: ما وراء البيانات في بيئة العنكبوتية العالمية: يمهد هذا الفصل للدراسة حيث يتناول البعد النظري لموضوع ما وراء البيانات مع التركيز على ما وراء البيانات في بيئة العنكبوتية، وعلاقتها بمحركات البحث، وخطط ما وراء البيانات في بيئة العنكبوتية، وأدوات إنشاء وتحليل ما وراء البيانات.

الفصل الثاني: تحليل ما وراء البيانات في مقالات الدوريات الإلكترونية العربية في مجال المكتبات والمعلومات: يبدأ الفصل بعرض للدوريات الإلكترونية العربية محل الدراسة، وعينة المقالات التي خضعت لتحليل ما وراء البيانات بها، ثم تناول تحليل ما وراء البيانات في مقالات الدوريات الإلكترونية العربية من حيث مدى توافرها في مقالات هذه الدوريات، وأبرز خطط ما وراء البيانات المستخدمة في تسجيل هذه العناصر، والتوزيع العددي والنوعي لتيجان ما وراء البيانات، ومدى جودة القيم الواردة في هذه التيجان، ومدى صلتها بالمقالات الواردة بها. كما تعرض هذا الفصل إلى نظرة استكشافية لعناصر ما وراء البيانات في عينة من مقالات ثلاثة من الدوريات الإلكترونية الأجنبية المتخصصة في مجال المكتبات والمعلومات.

الفصل الثالث: قياس تأثير ما وراء البيانات على كفاءة استرجاع مقالات الدوريات الإلكترونية العربية في مجال المكتبات والمعلومات: يتناول هذا الفصل اختبار تأثير ما وراء البيانات على استرجاع عينة من مقالات الدوريات الإلكترونية العربية في مجال المكتبات والمعلومات من ثلاثة محركات بحث هي: جوجل Google، وياهو Yahoo،

وبينج Bing، وبشكل خاص مـا وراء البيانـات التـي تـم تكويـدها باسـتخدام بـؤرة دبلـن، أو باستخدام مفردات سكيما مع إطار وصـف المصـدر في الخصـائص RDFa، بالإضـافة إلى ذلـك حاول هذا الفصل التحقق من إمكانية استثمار ما وراء البيانات في صفحات العنكبوتية بصفة عامة، والواردة في مقالات الدوريات الإلكترونية العربية في إعداد تسجيلات ببليوجرافية وفقـا لشكل الاتصال مارك 21 MARC21.

وتختم الدراسة بخاتمة تشتمل على النتائج والتوصيات، جاءت بعدها قائمة المصادر.

المؤلفة

الفصل الأول
ما وراء البيانات
في بيئة العنكبوتية العالمية

ويشتمل هذا الفصل على النقاط التالية:

- ✍ مصطلح ما وراء البيانات.
- ✍ تعريف ما وراء البيانات.
- ✍ أنواع ووظائف ما وراء البيانات.
- ✍ ما وراء البيانات والعنكبوتية العالمية.

الفصل الأول

ما وراء البيانات في بيئة العنكبوتية العالمية

تمهيد

يعد موضوع ما وراء البيانات من الموضوعات المتعددة الأوجه والمتشعبة الأطراف؛ ويرجع ذلك لاختلاف المجتمعات التي تتناوله والسياق الذي يستخدم فيه. فمن حيث التعريف، تتعدد التعريفات الخاصة بما وراء البيانات؛ فلا يوجد تعريف جامع وشامل لها.[1] كما تختلف دلالة ما وراء البيانات بين التخصصات والمجموعات التي تفيد منها سواء كانوا من المبرمجين ومصممي المواقع، أو أخصائيي الحاسبات الآلية، أو أخصائيي الاتصالات، أو أخصائيي المكتبات والمعلومات.[2] كما تختلف الرؤية حول الوظائف التي تقوم بها ما وراء البيانات وفقا للبيئات التي تعمل فيها سواء كانت مؤسسات المعلومات

---

(1) الجندي،محمود عبد الكريم. الاتجاهات الحديثة في الميتاداتا: مراجعة علمية للإنتاج الفكري. الاتجاهات الحديثة في مجال المكتبات والمعلومات. مج 14، ع 27. (يناير 2007). ص 93.

(2) المصدر السابق، ص 84.

أو قواعد البيانات أو الإنترنت.

ويحاول هذا الفصل توضيح الجوانب المختلفة المتعلقة بما وراء البيانات وخاصة في بيئة العنكبوتية العالمية وعلاقتها بمحركات البحث.

1 مصطلح ما وراء البيانات:

1-1 تاريخ المصطلح:

ينتمي مصطلح ما وراء البيانات في الأصل إلى علم الحاسب الآلي، وتتباين المعلومات حول أول من وضع هذا المصطلح؛ حيث تمت الإشارة إلى أن أول من وضع هذا المصطلح هو جاك إي مايرز Jack E. Myers في أواخر الستينات من القرن العشرين الميلادي وذلك في سياق "إدارة قواعد البيانات"[1] بينما تشير مصادر أخرى إلى أن أول من صاغ هذا المصطلح هو فيليب باجلي Philip Bagley في عام 1968 في كتابه "توسيع مفاهيم لغة البرمجة Extension of Programming Language Concepts" حيث كان من الواضح استخدامه للمصطلح بالمعنى التقليدي وهو "ما وراء البيانات المهيكلة Structural Metadata" بمعنى "بيانات عن حاويات البيانات Containers of Data" وذلك أكثر من المعنى الآخر المتعارف عليه حاليا وهو بيانات عن البيانات، أو نوع البيانات الذي عادة ما يوجد في فهارس المكتبات[2]. وقد تزايد استخدام المصطلح في

(1) كابلن، برسيلا. أساسيات ما وراء البيانات لاختصاصيي المكتبات والمعلومات / ترجمة وتعليق هاشم فرحات. الرياض: مكتبة الملك فهد الوطنية، 2007. مطبوعات مكتبة الملك فهد الوطنية. السلسلة الأولى، 54. ص 30، وأبو النور، إيناس محمد فوزي. تقنيات البيانات الخلفية لوصف المصادر الإلكترونية واسترجاعها على شبكة الإنترنت: دراسة ميدانية مقارنة / إشراف مصطفي حسام الدين. أطروحة (ماجستير) - جامعة القاهرة، كلية الآداب، قسم المكتبات والوثائق والمعلومات. 2008. ص 33.

(2) Metadata [online]. In: Wikipedia, The Free Encyclopedia. Available at: https://en.wikipedia.org/wiki/Metadata . [Accessed 1/9/2015].

الثمانينات من القرن نفسه حيث ظهر بصورة مبتكرة في الإنتاج الفكري الصادر عن نظم إدارة قواعد البياناتData Base Management System DBMS وذلك بغرض وصف المعلومات التي تشتمل عليها قواعد البيانات. ونظرا لأن الحاسب الآلي كان يمثل المجال الذي تم وصف البيانات فيه فقد ارتبط مصطلح ما وراء البيانات بشكل أساسي بالبيئة الإلكترونية.[1]

وفي عام 1986 استخدم مصطلح ما وراء البيانات كعلامة تجارية لشركة " Metadata Company" إحدى الشركات التي تقدم برمجيات وخدمات متخصصة في الطب والرعاية الصحية[2]. كما أن أول استخدام لهذا المصطلح في سياق تبادل البيانات، كان في الموجز الإرشادي لدليل ناسا لتبادل الأشكال الذي صدر عام 1988[3]. ومع أوائل التسعينات من القرن العشرين أصبح مصطلح ما وراء البيانات يستخدم بمعنى البيانات أو المعلومات الضرورية التي تيسر الإفادة من ملفات الحاسب، وخاصة عند التعامل مع مجموعات البيانات المتخصصة في العلوم الطبيعية والعلوم الاجتماعية ومجموعات البيانات الجغرافية المكانية. ومن أول المواصفات التي وسمت نفسها بأنها مواصفة ما وراء بيانات هي معيار محتوى ما وراء البيانات الجغرافية المكانية الرقمية Content Standard for Digital Spatial Metadata (CSDSM) للجنة الفيدرالية للبيانات الجغرافية التي نشرت إصدارته الأولى في عام 1994[4].

ومع انتشار الإنترنت والعنكبوتية بدأ مصطلح ما وراء البيانات يستخدم في سياق وصف مصادر المعلومات الرقمية المتاحة على الإنترنت وبخاصة بعد أن تبين أن الملفات النصية تحتاج إلى ما وراء بيانات تساعد على البحث عنها أو على الأقل إدارتها وضبطها.

---

(1) الجندي، محمود عبد الكريم. مصدر سابق. ص 83.
(2) كابلن، بريسيلا. مصدر سابق. ص 31.
(3) أبو النور، إيناس محمد فوزي. مصدر سابق. ص 33.
(4) كابلن، بريسيلا. مصدر سابق. ص 30- 32.

وفي عام 1995 دخل مصطلح ما وراء البيانات إلى المصطلحات العلمية المتداولة في مجال المكتبات مع استحداث مجموعة عناصر بؤرة دبلن Dublin Core لما وراء البيانات ؛ حيث كان معيار بؤرة دبلن بمثابة منطقة تفاعل بين مجتمع المكتبات من ناحية، ومجتمع العنكبوتية من ناحية أخرى، وساهم في إمداد مجتمع المكتبات بأفكار ومصطلحات جديدة[1].

وبعد انعقاد الورشة الأولى لمبادرة بؤرة دبلن Dublin core أدرك أخصائيو المكتبات والمعلومات أنهم كانوا يقومون منذ قرون بإنشاء ما وراء البيانات وذلك في سياق عملية الفهرسة التي يمارسونها، فالعملية نفسها – أي إنشاء ما وراء البيانات - كانت موجودة منذ القدم منذ قيام الإنسان بتنظيم المعلومات إلا أن مصطلح ما وراء البيانات هو الجديد وبخاصة في بيئة الحاسبات والشبكات الحالية.[2]

ويمكن الإشارة إلى أوجه التشابه بين ما وراء البيانات والفهرسة فيما يلي:

- تتشابه الأغراض الأساسية لإنشاء ما وراء البيانات مع تلك المرتبطة بالفهرسة حيث يهدف كلا منهما إلى تقديم وصف للمصادر أو الكيانات بما يتيح إمكانية الوصول إليها.
- تتشابه إجراءات إنشاء كل من ما وراء البيانات والفهرسة فكلتاهما تركز على تسجيل الخصائص التي تتيح للمستفيد إمكانية التعرف على مصادر المعلومات التي تلبي احتياجاته. حيث يتم تسجيل الخصائص سواء للمصادر الإلكترونية أو التقليدية من حيث العنوان، والمنشئ، والتاريخ، والموضوع ... إلخ من الخصائص التي تميز كل كيان وذلك كافي ليدل على العلاقة الوطيدة بين إنشاء ما وراء البيانات والفهرسة.

بينما هناك من يرى أنه من غير الإنصاف أن نساوي بين ما وراء البيانات والفهرسة وذلك لأن ما وراء البيانات تركز بصورة أكبر على المصادر الإلكترونية للمعلومات والتي

---

(1) المصدر السابق. ص 32 - 33.

(2) المصدر السابق. ص 33 – 34 .

بالطبع تختلف في فهرستها وخصائصها عن المصادر التقليدية[1].

وعلى الرغم من أن علم المكتبات والمعلومات يمتلك مصطلح الفهرسة إلا أنه يستخدم أيضا مصطلح ما وراء البيانات؛ ذلك لأن المتخصصين في مجال الحاسب الآلي والشبكات هم من أوجدوا هذا المصطلح أولا ثم شاركهم فيه أخصائيو المكتبات الذين يسعون إلى الالتقاء مع التخصصات الأخرى المهتمة بتنظيم المعلومات والتفاعل معها[2].فمصطلح ما وراء البيانات لم يعد قاصرا على مجال المكتبات فقط، كما هو الحال في مصطلح الفهرسة وإنما يتم استخدامه بشكل كبير في مجالات مختلفة. يضاف إلى ذلك أن مصطلح ما وراء البيانات عند انتشار استخدامه لأول مرة في مجال المكتبات، كان يقصد به ما وراء البيانات المضمنة في تيجان Tags صفحات العنكبوتية Web Pages المكتوبة بلغة ترميز النص الفائق HTML والتي لا يمكن للمستفيد العادي أن يراها إلا بعد فتح شفرة مصدر Source Code الصحفة، ومن هنا جاءت التسمية ما وراء البيانات Metadata. وهذا لم يكن في يوم من الأيام أحد أنشطة عملية الفهرسة كما أعتاد عليها أخصائيو المكتبات والمعلومات.

2-1 دلالة المصطلح:

يتكون مصطلح ما وراء البيانات Metadata من عنصرين هما meta من اليونانية، وdata من اللاتينية[3]. ويشير قاموس Oxford إلى أن البادئة meta تستخدم في بداية أي كلمة لتعني "تغير" أو "تحول" في المظهر أو الصيغة أو الظروف كما في "metamorphosis"،

(1) تيلور، ألين ج. تنظيم المعلومات. جزء 1. / تأليف ألين ج. تيلور، دانيل ن. جودري ؛ ترجمة هاشم فرحات. الرياض: جامعة الملك سعود، النشر العملي والمطابع،2012. ص 182 – 183.

(2) بامفلح، فاتن . الميتاداتا وتنظيم مصادر المعلومات الإلكترونية في المكتبات. دراسات عربية في المكتبات وعلم المعلومات. مج 7، ع 3. (2002). ص 24 – 54.

(3) Gill, Tony. Metadata and the Web [online]. In: Introduction to Metadata. Online Edition, Version 3.0. Los Angeles: Getty Research Institute, 2008. Available at: http://www.getty.edu/research/publications/electronic_publications/intrometadata/metadata.html . [Accessed 25/11/2013].

أو يمكـن أن تعنـي "وراء"، أو "مـا وراء"، أو "بعـد"، أو "مـا بعـد" كـما في مـا وراء الطبيعـة "metaphysics"، أو يمكن أن تعني "عن "[1].

والبادئـة meta في سـياق مصـطلحات علـوم الحاسـب تقابـل كلمـة about في اللغـة الإنجليزية، وهي ما تقابـل بالعربيـة في أبسـط معانيهـا "عـن"[2]. أمـا العـنصر الثـاني في هـذا المصطلح هي الكلمة اللاحقة "data " وهـي صـيغة الجمـع مـن الكلمـة اللاتينيـة "dutum"، وهي تشير في استخدامها العام إلى "المعطيـات"، أو "البيانـات" في أي شـكل كانـت أرقامـا، أو كلمات، أو صور، .. إلخ .

وهـذا المصـطلح في مجملـه يعنـي حرفيـا "البيانـات عـن البيانـات"، وهـي عبـارة عـن معلومات أو بيانات عن مجموعة أخرى مـن البيانـات. ومـن الأمثلـة الشـائعة عليهـا بطاقـة فهرس المكتبة التي تحتوي على بيانات عن محتويات الكتاب ومكانه[3].

3-1 المقابلات العربية للمصطلح:

تعددت الترجمات الخاصة بهذا المصطلح في الإنتاج الفكري العربي، فقد أشـارت إحـدى المراجعات العلمية[4] عن موضوع ما وراء البيانات إلى أن بداية الإنتاج الفكري العربي في هـذا الموضوع ترجع إلى عام 2000.

وقد بلغ عدد الترجمات المقترحة لهذا المصطلح نحو خمس عشرة ترجمة ؛ هـي:مـا وراء البيانـات، ومـا بعـد البيانـات، وبيانـات البيانـات، و بيانـات الوصـف المدمجـة في صفحات الإنترنت، ووصف/ وصائف البيانات، وواصفات البيانات، والبيانات الفارقة، وما فوق البيانات، والبيانات الواصفة، والبيانات الخلفية، وبيانـات عـن البيانـات، والبيانـات

(1) أبو النور، إيناس محمد فوزي. مصدر سابق. ص 32.

(2) هاشم فرحات. مقدمة المترجم:عن القضية وأبعادها. في: كابلن، بريسيلا. مصدر سابق. ص 21.

(3) Gill, Tony. Op Cit.

(4) الجندي، محمود عبد الكريم. مصدر سابق. ص 83.

الوصفية[1]، والبيانات الفوقية[2]- هذا فضلا عن النقل الصوتي للمصطلح في لغته الأصلية الذي لم يخل من الاختلاف أيضا -، وميتاديتا، وميتاداتا.

2 تعريف ما وراء البيانات:

مع تزايد تبني استخدام مصطلح ما وراء البيانات بواسطة العديد من المجتمعات المختلفة، فإن التعريف الخاص بما وراء البيانات تنامى ليشمل تقريبا أي شيء يصف أي شيء آخر. والتعريف المختصر المعياري لما وراء البيانات هو أنها "بيانات عن البيانات"[3].

ويزخر الإنتاج الفكري بالعديد من التعريفات لما وراء البيانات، ولا يوجد تعريف جامع مانع لما وراء البيانات ولكن ينبغي على من يتعامل مع هذا المصطلح أن يعلم أن مدلوله قد يختلف باختلاف المجتمع الذي يستخدمه، والسياق الذي يستخدم فيه[4]، وفيما يلي عرضا لعدد من التعريفات لهذا المصطلح:

يعرف قاموس المكتبات والمعلومات المتاح على الخط المباشر ODLIS ما وراء البيانات بأنها: حرفيا هي بيانات عن البيانات، وهي بيانات مهيكلة تصف مصادر المعلومات / أو كيانات Objects لتحقيق مجموعة متنوعة من الأهداف. وعلى الرغم من

(1) كما وردت في دراسة: عباس، بشار. البيانات الوصفية =Metadata: أدوات التصنيف والبحث عن المعلومات في العصر الرقمي. العربية 3000. ع 33 (2008). ص ص 34 – 37.

(2) كما وردت في دراسة: الزهيري، طلال ناظم. البيانات الفوقية للمواقع الحكومية العراقية على الإنترنت وتأثيرها في آلية تكشيفها من قبل محركات البحث [على الخط المباشر]. المجلة العراقية لتكنولوجيا المعلومات. مج 1،ع2 (2008). متاح على: http://www.slideshare.net/drtalal/ss-8472149.

(3) Gill, Tony. op.cit. Available at: http://www.getty.edu/research/publications/electronic_publications/intrometadata/metadata.html . [Accessed 25/11/2013].

(4) كابلن، بريسيلا. مصدر سابق. ص 36.

أن الفهرسة بقواعد الفهرسة الأنجلو أمريكية / شكل مارك AACR2/MARC تعد رسميا ما وراء بيانات، إلا أن مصطلح ما وراء البيانات يستخدم بصفة عامة في مجال المكتبات والمعلومات للخطط غير التقليدية مثل مجموعة عناصر بؤرة دبلن Dublin Core Metadata Element Set، وقطاعات بؤرة جمعية المصادر المرئية VRA Core Categories، والوصف الأرشيفي المرمز (Encoded Archival Description (EAD [1].

كما قدمت الموسوعة العالمية لعلم المكتبات والمعلومات في عام 2004 تعريفا لما وراء البيانات بأنها "بيانات مهيكلة تستخدم للإيجاد، والوصول، والاستخدام، والإدارة لمصادر المعلومات وبصفة خاصة في البيئة الإلكترونية "[2].

فيما قدم عدد من المؤسسات والمنظمات تعريفات لما وراء البيانات؛ فالاتحاد الدولي لجمعيات المكتبات ومؤسساتها – إفلا IFLA يعرفها بأنها " أية بيانات تساعد على التثبت من هوية المصادر الإلكترونية الشبكية، والتعريف بها ووصفها وتحديد أماكنها"[3] كما قدمت جمعية الناشرين الأمريكيين تعريفا ينص على أنها "معلومات تصف المحتوى، وأفضل مثال واقعي على ذلك هو الفهرس البطاقي في المكتبة، أو المدخل في الفهرس المطبوع في الكتاب، أو المعلومة في الكشاف المتاح على الخط المباشر[4]، كما عرفها مكتب شبكات المكتبات والمعلومات البريطانية بأنها تعني البيانات المهيكلة أو المنظمة عن المصادر الرقمية وغير الرقمية بحيث يمكن استخدام هذه البيانات في دعم مجموعة كبيرة من العمليات، منها – على سبيل المثال - وصف هذه المصادر واستكشافها، وإدارتها (بما في ذلك إدارة الحقوق)، هذا فضلا عن حفظها حفظا طويل الأمد[5].

---

(1) Reitz, Joan M. ODLIS [online]: Online Dictionary for Library and Information Science. Available at: http://www.abc-clio.com/ODLIS/odlis_m.aspx. [Accessed 29/1/2016].

(2) أبو النور، إيناس محمد فوزي. مصدر سابق. ص 45.

(3) كابلن، بريسيلا. مصدر سابق. ص 33.

(4) المصدر السابق. ص 35.

(5) المصدر السابق. ص 38 - 39.

كما يمكن تعريف ما وراء البيانات بأنها توفر معلومات عن محتوى مادة بعينها. على سبيل المثال، قد تشتمل صورة ما على ما وراء بيانات تصف حجم الصورة، وعمق الألوان بها، ودقة وضوح الصورة، وتاريخ إنشاءها، وغيرها من البيانات. وقد تشتمل ما وراء البيانات للوثائق النصية على معلومات عن مؤلفها، تاريخ كتابتها، وملخصها. وغالبا ما تتضمن صفحات العنكبوتية ما وراء البيانات في شكل تيجان الميتا Meta Tags. وعادة ما يتم استخدام كل من تاج الوصف Description وتاج الكلمات المفتاحية Keywords لوصف محتوى صفحات العنكبوتيةWeb Pages. وتستخدم معظم محركات البحث هذه البيانات[1].

إلا أن اتحاد العنكبوتية العالمية W3C قد تبنى تعريفا مقيدا لما وراء البيانات حيث عرفها بأنها "معلومات مفهومة آليا للعنكبوتية" وبالتالي فقد قصر هذه المعلومات على ما يوجد على العنكبوتية فقط، وبناء على ذلك فقد استبعد هذا التعريف ما وراء البيانات في مختلف أنواع الفهارس التي تقوم بإنشائها المكتبات ومراكز المعلومات أو التي يوفرها الناشرون، بل ويتعدى أيضا هذا التهميش إلى وصف المصادر التي يمكن الوصول إليها عن طريق بروتوكولات الإنترنت الأخرى التي لا تعتمد على العنكبوتية مثل FTP[2].

كما أشار عدد من الباحثين إلى تعريفات مختلفة لما وراء البيانات حيث يرى تيم بيرنرز لي – مخترع العنكبوتية - أن ما وراء البيانات هي معلومات مفهومة آليا عن المصادر المتاحة على العنكبوتية أو عن أشياء أخرى[3]. أما كوللر Koehler فيرى أن ما وراء البيانات هي

---

(1) Metadata Definition .In: TechTerms. Available at: http://techterms.com/definition/metadata . [Accessed 1/2/2016].

(2) عكوش، نبيل، غانم، نذير، كوكة، سليمة . ما وراء البيانات ومعاييرها في بيئة المكتبات. في: المؤتمر الدولي لتكنولوجيا المعلومات الرقمية: الاتجاهات الحديثة في تكنولوجيا المعلومات. (عمان:جامعة الزرقاء: 9 – 11 أكتوبر 2012) . متاح على: http://digital.jilwan.com/digital2012/download2012.php?f=jalsa1/1_4.pdf [تاريخ الزيارة 2016/1/30].

(3) Berners-Lee, Tim. Axioms of Web Architecture: Metadata [online]. 1997. Available at: https://www.w3.org/DesignIssues/Metadata.html. [Accessed 1/2/2016].

مصطلح جديد سك حديثا للتعبير عن مفهوم قديم في علم المكتبات، ويهدف إلى إيجاد وتسجيل بيانات عن المصادر الإلكترونية للمعلومات بحيث تسهل من سرعة استدعائها وتداولها ونقلها[1].

كما يشير إيليت Ellett إلى أن ما وراء البيانات ما هي إلا بيانات تصف سمات وخصائص مصادر المعلومات وتوضح علاقاتها وتساعد على الوصول إليها أو اكتشافها وإدارتها واستخدامها بفعالية[2]

ومما تجدر الإشارة إليه أن تعريفات ما وراء البيانات تبدو حتى بين اختصاصي ما وراء البيانات أنفسهم معقدة ومدعاة للخلط، ولعل هذا يرجع – من ناحية – إلى الطبيعة متعددة الأوجه لهذا الموضوع، كما يرجع من ناحية أخرى إلى التعريف المفرط في الاتساع والانتشار للمصطلح وهو "بيانات عن البيانات"[3].

ويتضح من التعريفات السابقة أنها في معظمها تتراوح ما بين التعريف في أوسع مستوياته من حيث أن ما وراء البيانات هي "بيانات عن البيانات"، انطلاقا من فرضية أن أي مصدر معلومات أو كيان هو شكل من أشكال البيانات، وبالتالي فإن وصف خصائص هذا المصدر وسماته ومحتوياته ستكون عبارة عن بيانات ؛ ومن ثم فهي بيانات عن البيانات إلى تعريفات أكثر تعقيدا وأطول صياغة، لكن القاسم المشترك بينها جميعا هو

---

(1) Koehler, Wallace. A Call to Action: What Every Searcher Should Know – And Do About Domain Names, Standards, and metadata. Searcher. Vol. 10, No. 9, (Oct.2002), p 22.6.

نقلا عن: الجندي، محمود عبد الكريم. مصدر سابق. ص 94.

(2) Ellett, Robert O. Internet Search Engines giving you Garbage [online]. Available at: http://scis.nova.edu/ellettro.

نقلا عن: الجندي، محمود عبد الكريم. مصدر سابق. ص 94.

(3) تيلور، ألين ج. مصدر سابق. ص 135 – 136.

أن ما وراء البيانات هي معلومات مهيكلة أو بنائية تصف خصائص مصادر المعلومات – بصرف النظر عن كونها رقمية أو غير رقمية وإن كانت في أغلبها تركز على المصادر الرقمية-، بغرض التعريف بها واكتشافها واختيارها واستخدامها وإتاحة الوصول إليها وإدارتها[1].

كما يلاحظ على هذه التعريفات أنها تركز بصفة أساسية على مصادر المعلومات باعتبارها الكيان الأساسي الذي تصفه ما وراء البيانات، ولكن من وجهة نظر الباحثة يمكن أن يتسع هذا المفهوم ليشمل وصف أي كيان Object (سواء كان مصدر معلومات، أو حدث، أو شخص، أو مفهوم، أو فكرة) وبخاصة في إطار التطورات التي تشهدها الإنترنت والعنكبوتية الآن والمتمثلة في الاتجاه نحو العنكبوتية الدلالية Semantic Web وهذا يتفق مع التعريف الذي قدمه تيم بيرنرز لي - كما ورد سابقا - والذي أشار فيه إلى أن ما وراء البيانات هي " معلومات مفهومة آليا عن المصادر المتاحة على العنكبوتية أو عن أشياء أخرى"، وقد فسر ذلك بقوله "في المستقبل، عندما تكون لغات ما وراء البيانات والمحركات أكثر تطورا، فيمكنها أن تشكل قاعدة قوية لعنكبوتية من المعلومات المفهومة آليا عن أي شيء: عن الأشخاص، أو الأشياء، أو المفاهيم، أو الأفكار"[2].

كما يؤكد هذه الرؤية تعريف جاين جرينبرج Jane Greenberg، والذي جاء فيه أن ما وراء البيانات هي "بيانات مهيكلة عن كيان ما، وتدعم هذه البيانات الوظائف المرتبطة بهذا الكيان". فالكيان كما عرفته المنظمة الدولية للتقييس / اللجنة الدولية للتقنية الإلكترونية International Organization for Standardization/ International Electrotechnical Commission (ISO / IEC) هو : أي جزء من العالم يمكن إدراكه أو تخيله. كما أن هناك تعريف آخر يرى أن الكيان يتخطى النصوص، والصور، والوسائط المتعددة ليشمل الكيانات الحية living entities. ووفقا لذلك فإن الكيان في هذا التعريف

(1) نفس المصدر السابق والصفحات.

(2) Berners-Lee, Tim . Op Cit.

يمتد ليشمل المصادر المادية والرقمية مثل المنفردات، والصحف، والصور وكذلك الأنشطة، والأحداث، والأشخاص، والبنى، والعلاقات، والبرامج[1].

وفي هذا الإطار تضع هذه الدراسة تعريفا عاما لما وراء البيانات بأنها "بيانات مهيكلة عن أي كيان (سواء كان مصدر معلومات – رقمي أو غير رقمي -، شخص، هيئة، خدمة...إلخ)، وذلك لتحقيق أهداف متنوعة مثل التعريف والاسترجاع والاستخدام والإدارة. ويقصد ببيانات مهيكلة في هذا التعريف أنها ترد في صيغة أو ترميز أو خطة ما.

أما التعريف الإجرائي لما وراء البيانات والذي تتخذه هذه الدراسة فهو: هي بيانات مهيكلة تصف صفحات العنكبوتية المكتوبة بلغة ترميز النص الفائق HTML بهدف تحديد الهوية والاسترجاع والإدارة (ويقصد بالإدارة هنا على سبيل المثال إمكانية التحكم في الصفحة أو الموقع من حيث إتاحة تكشيفه أو عدم تكشيفه من قبل محركات البحث، أو إتاحة الاستخدام والوصول لأفراد بعينهم).

3 أنواع ووظائف ما وراء البيانات:

يمكن التمييز بين أنواع ما وراء البيانات وتقسيمها وفقا لأكثر من أساس ولكن أكثرها شيوعا هو التقسيم وفقا لوظيفة ما وراء البيانات أو الغرض منها. ووفقا لهذا الأساس يمكن تقسيم ما وراء البيانات بصفة أساسية إلى ثلاث فئات واسعة هي: ما وراء البيانات الوصفية Descriptive Metadata، وما وراء البيانات البنائية Structural Metadata، وما وراء البيانات الإدارية Administrative Metadata [2].

(1) Greenberg, Jane. Metadata and the World Wide Web [online]. In: Encyclopedia of Library and Information Science. Vol. 3. 2nd ed. p1876. Available at: https://books.google.com.eg/books?id=Sqr-_3FBYiYC&printsec =frontcover&hl=ar#v=onepage&q&f=false . [Accessed 8/2/2016]..

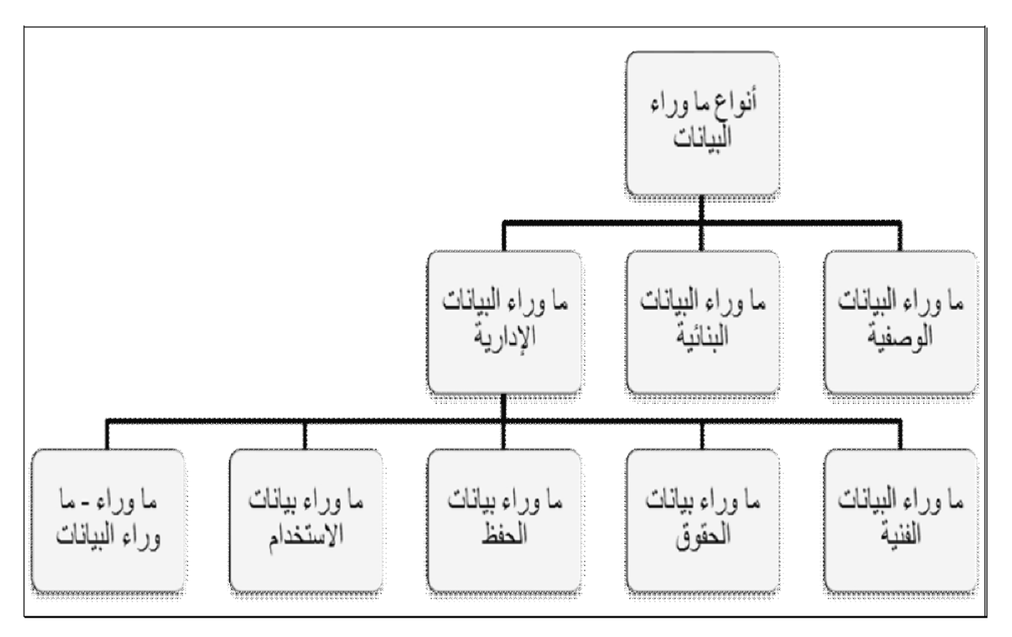
(2) تيلور، ألين ج. مصدر سابق. ص 138.

وتجدر الإشارة إلى أن هذه التقسيمات ما هي إلا تقسيمات نظرية أما من الناحية العملية، فهي قد تتداخل مع بعضها البعض[1]، فالحدود بينها غير واضحة أو قاطعة ؛ فعلى سبيل المثال يمكن تصنيف بعض عناصر وراء البيانات في أكثر من فئة واحدة ؛ مثال ذلك المحدد الفريد Unique Identifier قد يعامل على أنه ضمن ما وراء البيانات الوصفية وكذلك ينتمي لما وراء البيانات الإدارية أو البنائية. كما أن ما يعتبره أحد المؤلفين ما وراء بيانات فنية، قد يعتبره مؤلف آخر ما وراء بيانات بنائية أو إدارية، كما أن الأصل والأساس أن جميع ما وراء البيانات هي وصفية حيث أنها تصف شيئا ما.

ويضاف إلى هذه الفئات الثلاث الرئيسية فئات أخرى إضافية ترد جميعها تحت ما وراء البيانات الإدارية وتشمل: ما وراء البيانات الفنية Technical Metadata، وما وراء بيانات الحقوق Rights Metadata، وما وراء بيانات الحفظ Preservation metadata، وما وراء بيانات الاستخدام Use Metadata هذا فضلا عن ما وراء البيانات المتعدية أو ما وراء ما وراء البيانات Meta-Metadata[2] التي تصف ما وراء البيانات نفسها. ويوضح الشكل رقم (1-1) أنواع ما وراء البيانات وفقا لوظيفتها:

(1) عكوش، نبيل، غانم، نذير، كوكة، سليمة. مصدر سابق.

(2) تيلور، ألين. ج. مصدر سابق. ص 138، وإيناس أبو النور. مصدر سابق. ص 57، وعبد الهادي، محمد فتحي. الميتاداتا: أسسها النظرية وتطبيقاتها العملية / محمد فتحي عبد الهادي، خالد عبد الفتاح. القاهرة: الدار المصرية اللبنانية، 2013. ص 53.



الشكل رقم (1-1) أنواع ما وراء البيانات وفقا للوظيفة

1-3 ما وراء البيانات الوصفية:

هي البيانات التي تصف الخصائص المحدِدة أو المميزة لمصدر المعلومات ومعها تحليل محتوياته الفكرية وتشمل:

- البيانات التي تحدد هوية مصدر المعلومات مثل العنوان، المؤلف، تاريخ النشر.
- بيانات التنظيم الفكري مثل الضبط الاستنادي والعلاقات بين الكيانات.
- بيانات الوصول الفكري مثل رؤوس الموضوعات.

وعموما فإن ما وراء البيانات الوصفية مثل البيانات الببليوجرافية التقليدية، تحدد هوية المصدر وتصف خصائصه، وتصور علاقاته مع المصادر الأخرى، وتحدد التقييم للمصدر.[1]

(1) عبد الهادي، محمد فتحي. مصدر سابق. ص 55 - 56.

وتدعم ما وراء البيانات الوصفية عمليات استكشاف مصادر المعلومات والعثور عليها، واختيارها (أي تحديد ما إذا كان مصدر ما يلبي احتياجات بعينها)، كما تستخدم في عمليات الاقتناء، والتقييم، والربط (أي التعبير عن العلاقات بين المصدر الموصوف من جهة وغيره من المصادر)، ومدى القابلية للاستخدام (مثل تحديد الأجهزة والبرمجيات اللازمة للاستخدام)[1].

ومن أمثلة ما وراء البيانات الوصفية تسجيلات الفهرسة، ووسائل الإيجاد، والكشافات المتخصصة[2]، وتيجان الميتا في لغة ترميز النص الفائق HTML Meta Tags.

2-3 ما وراء البيانات البنائية:

هي البيانات التي تشير إلى هيكل أو بناء مصدر المعلومات، أو مجموعة من المصادر فهي تعبر عن التنظيم الداخلي للمصدر والروابط والعلاقات بين أجزائه [3]، وتضمن أن مصدر المعلومات يعمل بشكل صحيح، ويمكن للمستفيد أن يستخدمه ويتصفحه. وقد يشار إليها في بعض الأحيان بما وراء بيانات العرض Display Metadata، كما أنه في أحيان أخرى تدمج خطأ مع ما وراء البيانات الفنية [4].

ويمكن تشبيه وظيفة ما وراء البيانات البنائية بالمادة الصمغية التي تعمل على تماسك مصادر المعلومات الرقمية بعضها مع بعض. فالكتاب مثلا قد يحتوي على عدة فصول، وكل فصل يتألف من عدة صفحات، ومن ثم فإن كل صفحة من هذه الصفحات يمثلها ملف رقمي منفصل؛ وهنا تكمن وظيفة ما وراء البيانات البنائية في تسجيل الروابط بين الملفات والصفحات المادية من ناحية، وبين الصفحات والفصول من ناحية ثانية، وبين

---

(1) كابلن، بريسيلا. مصدر سابق. ص 39 - 42.
(2) عبد الهادي، محمد فتحي. مصدر سابق. ص 53.
(3) أبو النور، إيناس فوزي. مصدر سابق. ص 56 - 57.
(4) تيلور، ألين ج. مصدر سابق. ص 154 - 155.

الفصول والكتاب كله من ناحية ثالثة. وبهذا المنطق تستخدم برمجيات العرض ما وراء البيانات البنائية في عرض قوائم المحتويات، وتوظيفها في عمليات الانتقال المباشرة إلى فصل معين، أو تقليب الصفحات للأمام أو للخلف. كما أن ما وراء البيانات البنائية تربط مكونات الوسائط المتعددة بعضها ببعض، كأن تعمل على ربط الصوت مع النص المكتوب، وضبط تزامن صوت المعلق مع النص المكتوب، وتشترك ما وراء البيانات البنائية مع ما وراء البيانات الوصفية التي تخدم قابلية الاستخدام في بعض الجوانب وإن كان الفارق بينهما في أن ما وراء البيانات الوصفية الخاصة بمدى قابلية الاستخدام موجهة أساسا للبشر، بينما تستخدم ما وراء البيانات البنائية في عمليات المعالجة التي تقوم بها الحاسبات[1].

فضلا عن ذلك فإن ما وراء البيانات البنائية تتيح إمكانية الربط بين أشكال التمثيل المختلفة للمحتوى الفكري نفسه، حيث يمكن لكتاب ما أن يتوافر بأكثر من صيغة واحدة للعرض، وهنا تبرز أهمية ما وراء البيانات في ربطها هذه الصيغ بعضها مع بعض، ومن ثم يمكن للمستفيد أن يختار أسلوب العرض الذي يناسبه، كما أنها تلعب دورا في إدارة المصادر وحفظها؛ فالمستودعات الرقمية التي تتولى اختزان هذه المصادر تحتاج إلى معرفة الملفات التي تتكون منها مصادر المعلومات حتى تتمكن من تنفيذ أية إجراءات عليها مثل عمليات الإضافة Ingestion، وكتابة التقارير، والبث أو التوزيع[2].

3-3 ما وراء البيانات الإدارية:

هي البيانات التي تستخدم في إدارة مصادر المعلومات[3]، وتشمل وظائف الإدارة تتبع المصدر عبر مختلف مراحل معالجته، وضبط إتاحته أو الوصول إليه، والتحقق من المسؤوليات المتصلة به، ومنح تصاريح استخدامه أو الإفادة منه. ويمكن القول أن ما وراء البيانات الوصفية توجه أساسا لخدمة المستفيد الفعلي أو المحتمل من المصدر، بينما توجه ما

(1) كابلن، بريسيلا. مصدر سابق. ص 44- 45.
(2) المصدر السابق. ص 438 - 439.
(3) عبد الهادي، محمد فتحي. مصدر سابق. ص 53.

وراء البيانات الإدارية لخدمة صاحب هذا المصدر أو القائم على أمره[1].وتدعم ما وراء البيانات الإدارية وظائف الإدارة واتخاذ القرار وحفظ السجلات وتقدم معلومات عن متطلبات الاختزان وعمليات الانتقال migration للمصادر، كما تساعد في مراقبة واستنساخ ورقمنة مصادر المعلومات.

وتشمل معلومات مثل:

- معلومات الاقتناء مثل كيف ومتى أنشئ المصدر، ومعلومات إدارية عن المصدر التناظري الذي اشتق منه المصدر الرقمي.
- معلومات الملكية، والحقوق، والسماح، والاستنساخ.
- متطلبات الإتاحة القانونية ومن له حق استخدام المادة.
- معلومات مكان المصدر مثل الموقع الموحد للمصدر URL أو رقم الطلب.
- معلومات الاستخدام مثل المستفيدين، وإعادة استخدام المحتوى.
- معلومات الحفظ مثل معلومات توثيق الحالة المادية، وتوثيق الأنشطة أو الأعمال اللازمة للحفظ[2].

وتنقسم ما وراء البيانات الإدارية إلى فئات فرعية يصعب حصرها وإن كان من أهمها[3]:

ما وراء البيانات الفنية Technical Metadata، وما وراء بيانات الحقوق Rights Metadata، وما وراء بيانات الحفظ Preservation metadata، وما وراء بيانات الاستخدام Use Metadata، وما وراء – ما وراء البيانات Meta – Metadata. وفيما يلي نبذة عن كل نوع من هذه الأنواع:

(1) كابلن، بريسيلا. مصدر سابق. ص 416.

(2) عبد الهادي، محمد فتحي. مصدر سابق. ص 56 – 57.

(3) كابلن، بريسيلا. مصدر سابق. ص 44.

أ- ما وراء البيانات الفنية Technical Metadata:

تستخدم في توثيق البرامج والمكونات المادية وتعمل على توثيق وتأمين البيانات مثل كلمات السر ومفاتيح التشفير وتتبع زمن استجابة النظام(1). كما أنها تعد أساسية لإجراءات حفظ المصدر للاستخدام المستقبلي(2).

ب- ما وراء بيانات الحقوق Rights Metadata

هي بيانات عن من يملك حق الوصول لمصادر المعلومات واستخدامها، ولأي الأغراض. وهي تتعلق بحقوق الملكية الفكرية للمنشئين والاتفاقات القانونية التي تسمح للمستفيدين بالوصول لهذه المعلومات. وتشمل عناصر ما وراء بيانات الحقوق على: فئات الوصول أو الإتاحة، والمعرفات Identifiers، وبيانات حقوق الطبع، والشروط، وفترات الإتاحة، ومعلومات الاستخدام، واختيارات الدفع(3).

ج- ما وراء بيانات الحفظ Preservation Metadata:

هي البيانات المتصلة بإدارة عملية حفظ مصادر المعلومات، وهي لازمة لضمان استمرار استخدام وإتاحة مصادر المعلومات في المستقبل(4)، وذلك في ظل التغيرات السريعة في التقنية والبرمجيات والعتاد. ومن أهم عناصر هذه الفئة من ما وراء البيانات: نوع البنية، ووصف الملف، والحجم، والخصائص المميزة، وبنية البرمجيات والعتاد، ومعلومات المصدر، وتاريخ الكيان، وتاريخ التحويل أو النقل(5).

---

(1) الجندي، محمود عبد الكريم. مصدر سابق. ص 97.

(2) تيلور، ألين ج. مصدر سابق. ص 150.

(3) عبد الهادي، محمد فتحي. مصدر سابق. ص 58.

(4) عبد الجواد، سامح زينهم. خطط ومعايير ميتاداتا الحفظ الرقمي: دراسة تحليلية.1. مجلة المكتبات والمعلومات العربية. س 30، ع 1. ص 165.

(5) تيلور، ألين ج. مصدر سابق. ص 152.

د- ما وراء بيانات الاستخدام Use Metadata:

هي ما وراء البيانات المتصلة بمستوى ونوع استخدام مصادر المعلومات مثل تعقب الاستخدام والمستخدمين[1].

هـ - ما وراء - ما وراء البيانات Meta – Metadata:

هي بيانات عن ما وراء البيانات. ويمكن أن تشمل بيانات عن منشئ ما وراء البيانات، ووقت إنشاءها، وكيفية انشاءها[2].

وتضيف الباحثة إلى التقسيم السابق المبني على الوظيفة، تقسيما آخرا يتناسب مع طبيعة هذه الدراسة، وهو تقسيم يعتمد على التمييز بين ما وراء البيانات وفقا للبيئة التي تعمل فيها أو توجد بها، ووفقا لذلك يمكن تقسيم ما وراء البيانات إلى:

1- ما وراء بيانات العنكبوتية: هي ما وراء البيانات التي تتوافر في بيئة العنكبوتية العالمية، والتي تأتي في أغلب الأحوال مصاحبة للمصادر المتاحة على العنكبوتية في تيجان الميتا Meta Tags التي ترد في الصفحات المكتوبة بلغة ترميز النص الفائق HTML – أو في لغة ترميز النص الفائق الموسعة XHTML التي تشكل أحد مشتقاتها -، والتي تم التفكير فيها في بداية التسعينيات كأحد الوسائل التي يمكن لمحركات البحث أن تستعين بها لتنظيم المعلومات على العنكبوتية، وهذه هي الفئة التي معها ذاع مصطلح ما وراء البيانات. كما يوصف هذا النوع من ما وراء البيانات عادة بالحواشي أو التعليقات annotations حيث أنها بمثابة ملاحظات ترتبط بالمصادر[3].

---

(1) عبد الجواد، سامح زينهم.. مصدر سابق. ص 165.

(2) Greenberg, Jane. Op. Cit. p 1881

(3) Kinsella, Sheila. Augmenting Social Media Items with Metadata using Related Web Content [online]/ Supervisor John Breslin. Digital Enterprise Research Institiute, National University of Ireland, Galway, January 2012 (Ph.D.). p 45 – 46. Available at: https://aran.library.nuigalway.ie/xmlui/bitstream/handle/10379/2674/paper.pdf?sequence=1 .[Accessed 17/2/2016].

وتركز الدراسة الحالية على هذه الفئة من ما وراء البيانات.

2- ما وراء البيانات الببليوجرافية: هي ما وراء البيانات التي توجد بشكل أساسي في فهارس المكتبات والمستودعات الرقمية، ومن أشهر أمثلتها تسجيلات الفهرسة المقروءة آليا بصيغة مارك 21 MARC 21. كما يمكن أن نطلق عليها ما وراء بيانات الفهارس والمستودعات الرقمية أو ما وراء بيانات مؤسسات المعلومات.

3- ما وراء بيانات نظم إدارة قواعد البيانات: هي ما وراء البيانات الخاصة بقواعد البيانات والتي شكلت البيئة الأولى التي ظهر بها مصطلح ما وراء البيانات قبل ظهور العنكبوتية، والتي تستخدم فيها ما وراء البيانات لتوثيق خصائص البيانات الواردة في هذه القواعد ؛ حيث تعمل على توفير بيانات عن كل عنصر من العناصر الواردة في قاعدة البيانات مثل اسم العنصر، ونوعه، والقيم الخاصة به. ويعرف هذا النوع أيضا باسم قاموس البيانات Data Dictionary [1].

4- ما وراء بيانات البرامج: هي ما وراء البيانات التي تنتجها البرامج والتطبيقات للملفات التي تنشئها وتعالجها، فعلى سبيل المثال لا الحصر برنامج مايكروسوفت وورد Microsoft Word يقوم بشكل تلقائي بإضافة البيانات التي تصف الملفات التي يُنشئها ومن أمثلة هذه البيانات العنوان، والمؤلف، وحجم الملف، وتاريخ الإنشاء والتعديل. ويمكن رؤية هذه البيانات كما يوضحها الشكل رقم (1-2) تحت عنصر الخصائص properties.

---

(1) Thakur, Dinesh. What is Metadata OR Data Dictionary? [online]. In: eComputer Notes. Available at: http://ecomputernotes.com/fundamental/what-is-a-database/metadata-or-data-dictionary .[Accessed 6/2/2016].



الشكل رقم (2-1) مثال على ما وراء البيانات في برنامج مايكروسوفت وورد

4 ما وراء البيانات والعنكبوتية العالمية:

ارتبطت ما وراء البيانات بالعنكبوتية العالمية ارتباطا وثيقا حيث انتشر استخدام هذا المصطلح عندما دعت الحاجة إلى التفكير في وسيلة لتنظيم المعلومات على العنكبوتية، وكانت ما وراء البيانات هي الوسيلة التي تم التوصل إليها. ومنذ ذلك الوقت ومع التطورات المختلفة التي شهدتها العنكبوتية اختلف استخدام ما وراء البيانات، وفيما يلي شرح لمراحل تطور العنكبوتية ودور ما وراء البيانات في كل مرحلة من هذه المراحل، وما وراء البيانات وعلاقاتها بأدوات تنظيم المعلومات على العنكبوتية وبشكل خاص محركات البحث.

1-4 ما وراء البيانات وأجيال العنكبوتية:

العنكبوتية العالمية هي وسيط عالمي لتبادل المعلومات يتيح للمستفيدين إمكانية القراءة والكتابة بواسطة الحاسبات الآلية المتصلة بالإنترنت. وهي بمثابة نظام من الوثائق المترابطة المتاحة عبر الإنترنت. ومع برنامج تصفح يمكن لأي فرد رؤية الصفحات التي

قد تحتوي على نصوص، أو صور، أو مقاطع فيديو إلى غير ذلك من الوسائط والانتقال بينها عبر الروابط الفائقة[1].

ومرت العنكبوتية بالعديد من التطورات منذ نشأتها. حيث بدأت بالعنكبوتية 1.0 Web 1.0 باعتبارها شبكة لربط الوثائق، ثم العنكبوتية 2.0 Web 2.0 كشبكة للروابط بين الأشخاص، ثم العنكبوتية 3.0 Web 3.0 والتي تسعى لأن تكون شبكة للروابط بين المعرفة أو البيانات، والعنكبوتية 4.0 Web 4.0 وهي شبكة الروابط الذكية والتي لا تزال فكرة قيد التطور ولا يوجد لها تعريف محدد حتى الآن. وتعرف أيضا بالعنكبوتية التكافلية أو التعايشية Symbiotic Web. وعلى الرغم من عدم وجود فكرة دقيقة عن العنكبوتية 4.0 وتقنياتها، لكن من الواضح أن العنكبوتية تتجه نحو استخدام الذكاء الاصطناعي لتصبح بمثابة شبكة ذكية.[2].

ويشكل كل جيل من هذه الأجيال أحد التطورات في حياة العنكبوتية بما يرفع من إمكاناتها، وبالتالي فإن ظهور جيل جديد من العنكبوتية لا يعني أنه يلغي الجيل الذي يسبقه وإنما يضيف إليه،. كما أن بعض هذه الأجيال تعبر عن أفكار لا تزال قيد التطبيق مثل العنكبوتية 3.0، والبعض الآخر يعبر عن تصورات أو أفكار مستقبلية مثل العنكبوتية 4.0. وفيما يلي شرح لأجيال العنكبوتية الأساسية وهي الأجيال الثلاثة وعلاقة ما وراء البيانات بكل جيل.

(1) Choudhury, Nupur. World Wide Web and Its Journey from Web 1.0 to Web 4.0.[online]. (IJCSIT) International Journal of Computer Science and Information Technologies. Vol. 5, No. 6. (2014). Available at: www.ijcsit.com/docs/Volume%205/vol5issue06/ijcsit20140506265.pdf . [Accessed 8/2/2016].

(2) Aghaei, Sareh, Nematbakhsh, Mohammad Ali and Farsani, Hadi Khosravi. Evolution of the World Wide Web from Web 1.0 to Web 4.0 [online]. International Journal of Web & Semantic Technology (IJWesT). Vol.3, No.1(January 2012). Available at: www.airccse.org/journal/ijwest/papers/3112ijwest01.pdf . [Accessed 8/2/2016]

1-1-4 ما وراء البيانات والعنكبوتية 1.0:

تعرف العنكبوتية 1.0 "Web 1.0" بالجيل الأول من العنكبوتية وتحدد بالفترة منذ بداية ظهور العنكبوتية على يد تيم بيرنرز لي عام 1989 حتى عام 2005. ووفقا لتيم بيرنرز لي فإن الجيل الأول من العنكبوتية هو مجرد شبكة قراءة فقط read-only web ؛ حيث تـوافر فيها قدر ضئيل جدا من التفاعل ففيها لا يمكن للمستفيدين التفاعل مع المواقع. وبعبارة أخرى فإن الشبكة في ذلك الوقت سمحت فقط بقراءة المعلومات وتصفحها والبحث عنها.[1]

ومع انتشار استخدام العنكبوتية وتوافر إمكانات الاتصال بها زاد حجم المعلومات المتاحة عليها بشكل متسارع، مما أدى إلى التفكير في وسيلة تسهل القدرة على استرجاع هذه المعلومات ومحاولة تنظيمها، وهنا ظهرت فكرة ما وراء البيانات كوسيلة لاكتشاف المصادر. واعتمدت فكرة ما وراء البيانات على إدراج عناصر لوصف الصفحات المتاحة على العنكبوتية من خلال تيجان Tags تسجل في جزء الرأس Head من صفحة مكتوبة بلغة ترميز النص الفائق HTML، وهذه التيجان لا يمكن للمستفيد النهائي رؤيتها إلا بعد فتح شفرة مصدر الصفحة Source Code، ومن هنا جاءت التسمية ما وراء البيانات.

ولم يمض وقت طويل منذ نشأة العنكبوتية في عام 1989 والبداية الفعلية لها كخدمة عامة على الإنترنت عام 1991[2] حتى ظهور معايير ما وراء البيانات، فقد شهدت السنوات الأولى من حياة العنكبوتية ظهور معايير ما وراء البيانات مثل تيجان الميتا HTML META Tags، ومجموعة عناصر بؤرة دبلن لما وراء البيانات Dublin Core Metadata Element Set (DCMES)، وإطار وصف المصدر[3] Resource

---

(1) Ibid.

(2) History of the World Wide Web [online]. In: Wikipedia: The Free Encyclopedia. Available At: https://en.wikipedia.org/wiki/History_of_the_ World_Wide_Web .Accessed [8/2/2016]..

(3) Gill, Tony. Op. Cit.

Description Framework (RDF). حيث ظهرت تيجان الميتا HTML META Tags في عام 1995 كما وردت في الإصدارة الثانية من لغة ترميز النص الفائق HTML 2.0 [1] والتي تعد أول إصدارة لهذه اللغة قصد بها أن تعامل معاملة المعيار[2]. كما ظهر معيار بؤرة دبلن في عام 1995، تلاه معيار إطار وصف المصدر RDF في عام 1999.

ووفقا لذلك فقد كان هناك توجها قويا نحو استخدام ما وراء البيانات، وشهدت بدايات عدد كبير من محركات البحث استخدام تيجان الميتا HTML META Tags بهدف تكشيف المواقع[3]، إلا أن استخدام ما وراء البيانات من قبل بعض الأفراد في خداع محركات البحث عن طريق إدراج ما وراء بيانات لا تمثل صفحاتهم أدى إلى تجاهل معظم محركات البحث لتيجان ما وراء البيانات[4].

2-1-4 ما وراء البيانات والعنكبوتية 2.0:

العنكبوتية 2.0 "Web 2.0" هي الجيل الثاني من العنكبوتية، ويصعب تحديد بداياتها بدقة[5]، ولكن يمكن الإشارة إلى ظهور هذا المصطلح رسميا في عام 2004 من قبل دايل دوجيرتيDale Dougherty والذي عرفها بأنها شبكة للقراءة والكتابة read-write web [6].

(1) Berners-Lee, T, Connolly, D. Hypertext Markup Language - 2.0 [online]. November 1995. Available at: http://tools.ietf.org/html/rfc1866. [Accessed 13/2/2016].

(2) HTML [online]. In: Wikipedia: The Free Encyclopedia. Available at: https://en.wikipedia.org/wiki/HTML . [Accessed 13/2/2016].

(3) عبد الهادي، محمد فتحي. مصدر سابق. ص 237.

(4) تيلور، ألين ج. مصدر سابق. ص 222.

(5) النشرتي، مؤمن سيد. محركات البحث الدلالية على الشبكة العنكبوتية: دراسة مسحية تحليلية لوضع مواصفات محركات بحث دلالية عربية / إشراف أسامة السيد محمود. أطروحة (ماجستير) - جامعة القاهرة. كلية الآداب. قسم المكتبات والوثائق وتقنية المعلومات، 2012. ص 100.

(6) Aghaei, Sareh, Nematbakhsh, Mohammad Ali and Farsani, Hadi Khosravi. Op Cit..

وهذا الجيل من العنكبوتية يعتمد بصفة أساسية على ما يقدمه المستفيدون. وتشمل العنكبوتية الثانية تقنيات وخدمات مثل المدونات وRSS والويكي Wikis، والتيجان Tags، والفلكسونومي Folksonomy [1]، كما تشتمل على الشبكات الاجتماعية مثل الفيس بوك Facebook، وماي سبايس MySpace، وتويتر Twitter، بالإضافة إلى مواقع تشارك الوسائط مثل اليوتيوب YouTube، وفليكر Flicker، وسليد شير Slideshare، وكذلك الشبكات الاجتماعية للمواقع المفضلة Social Bookmarking مثل ديلشيوس Delicious [2].

ومع ظهور هذا الجيل الجديد من العنكبوتية ظهرت الحاجة إلى استخدام ما وراء البيانات لتنظيم ما يتوافر عليها من مصادر ولتسهيل عمليات البحث والاسترجاع ولكنها ظهرت في أشكال مختلفة. حيث ظهرت ما وراء البيانات المنشأة من قبل المستفيدين User-Created Metadata [3] والبعض يطلق عليها ما وراء البيانات الاجتماعية Social Metadata والتي يمكن تعريفها بأنها معلومات إضافية عن مصدر المعلومات تأتي كنتيجة لمساهمات المستفيدين وأنشطتهم على الخط المباشر – مثل التوسيم Tagging، والتعليقات Comments، والمراجعات، والتقييمات، والتوصيات – والتي تساعد الأفراد على إيجاد، وفهم وتقييم المحتوى [4].

---

(1) Ibid.

(2) Patel, Karan. Incremental Journey for World Wide Web: Introduced with Web 1.0 to Recent Web 5.0 – A Survey Paper [online]. International Journal of Advanced Research in Computer Science and Software Engineering. Vol 3, No. 10 (October 2013). Available at: http://www.ijarcsse.com/docs/papers/Volume_3/10_October2013/V3I10-0149.pdf . [Accessed 8/2/2016].

(3) Gilliland, Anne J . Setting the Stage [online]. In: Introduction to Metadata. Op. Cit. Available at: http://www.getty.edu/research/publications/electronic_publications/intrometadata/setting.html .[Accessed 25/11/2013].

(4) Smith-Yoshimura, Karen, Shein, Cyndi. Social Metadata for Libraries, Archives and Museums [online]. part 1: Site Reviews. Dublin, Ohio:OCLC Research, 2011.p. 10. Available at: http://www.oclc.org/research/publications/library/2011/2011-02.pdf . [Accessed 9/2/2016].

ووفقا لذلك ينظر البعض إلى الفلكسونومي Folksonomy الذي يتم إنشاءه من خلال التوسيم Tagging في العديد من المواقع باعتباره مثال لما وراء البيانات المنشأة من قبل المستفيدين[1]. فعلى سبيل المثال يستخدم تويترTwitter ما يسمى هاشتاج Hashtag وهو عبارة عن رمز # تأتي بعده كلمات مفتاحية لتمييز الرسائل التي ترسل على هذا الموقع[2].

وفي السنوات الأخيرة، اتجهت شبكات التواصل الاجتماعي في ظل ما يتوافر عليها من كم ضخم من المعلومات وبشكل خاص ما يتشاركه المستفيدون من صفحات ومصادر متنوعة إلى استخدام ما وراء البيانات لوصف هذه المصادر؛ حيث ظهر ما يسمى بما وراء بيانات وسائل التواصل الاجتماعي Social Media Metadata التي تسمح بتوفير عناصر لوصف الصفحات والمصادر التي يتشاركها الأفراد على هذه الشبكات، والتحكم في طريقة ظهورها وعرضها عليها بما يجذب إليها المزيد من الزائرين[3]. فقد أعلن موقع الفيس بوك Facebook عن بروتوكول الرسم البياني Open Graph Protocol (og) الذي يعتمد على إضافة عناصر ما وراء البيانات لأي صفحة بما يجعلها تصبح كيانا أو عنصرا Object في الرسم البياني للفيس بوك Social Graph [4]، والذي يقصد به "الرسم الذي يصور العلاقات الشخصية بين مستخدمي الإنترنت" أو " الخريطة العالمية للجميع وكيفية ترابطهم معا"، حيث يكون في هذه الخريطة أو الرسم الكيانات بمثابة نقاط أو عقد Nodes والعلاقات فيما بينها عبارة عن خطوط تسمى حد Edge [5].

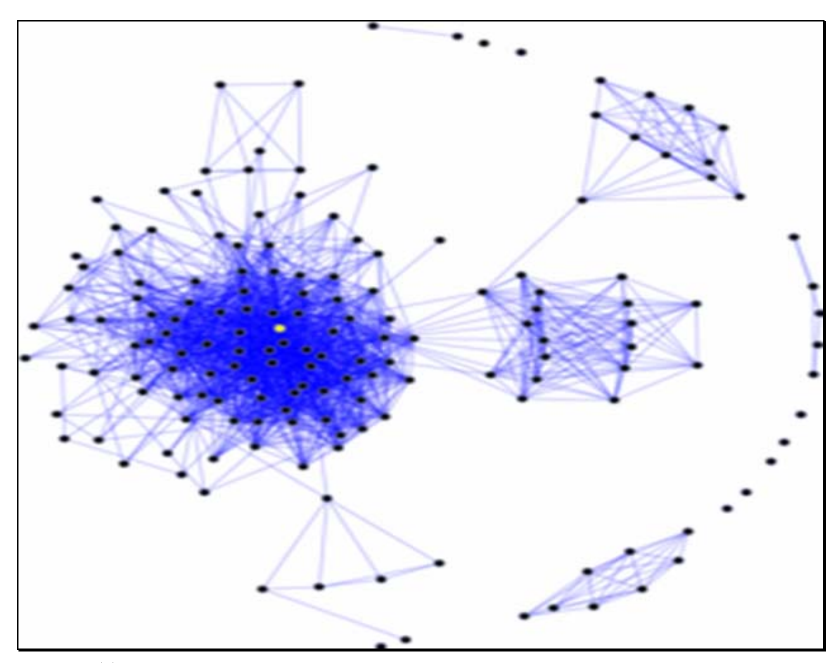
(1) Gilliland, Anne J .Op Cit.

(2) Using hashtags on Twitter. Help Center. Available at: https://support.twitter.com/articles/49309?lang=en . [Accessed 16/2/2016].

(3) Social Media Metadata: Facebook Open Graph and Twitter Cards [online]. Available at: https://www.websightdesign.com/services/internet-marketing/ search-engine-optimization/social-media-metadata . [Accessed 17/2/2016].

(4) The Open Graph protocol [online]. Available at: http://ogp.me/ . [Accessed 7/9/2014].

(5) Social graph [online]. In: Wikipedia: The Free Encyclopedia. Available at: https://en.wikipedia.org/wiki/Social_graph . [Accessed 16/2/2016].



الشكل رقم (1-3) الرسم البياني الاجتماعي Social Graph [1]

كما اتجه تويتر إلى استخدام ما يسمى ببطاقات تويتر Twitter Cards وهي عبارة عن عناصر لما وراء البيانات توضع في شفرة مصدر صفحات العنكبوتية التي ترد في الرسائل على تويتر[2]. وسيرد في جزء معايير ما وراء بيانات العنكبوتية تفصيلا لهذه المعايير.

3-1-4 ما وراء البيانات والعنكبوتية 3.0:

تعد العنكبوتية 3.0 "Web 3.0" الجيل الثالث من العنكبوتية، وتعرف بالعنكبوتية الدلالية Semantic Web. وترجع فكرة العنكبوتية الدلالية إلى عام 2001 مع ظهور

---

(1) Ibid .

(2) Twitter. Twitter Cards [online]. Available at: https://dev.twitter.com/cards/overview . [Accessed 16/2/2016].

المقالة المعنونة بـ "العنكبوتية الدلالية The Semantic Web " لمخترع العنكبوتية تيم بيرنرز لي وجيمس هينلدر James Hendler و أورا لاسيلا Ora Lassila والتي تم فيها نشر التعريف الأشهر للعنكبوتية الدلالية والذي ينص على أنها " ليست عنكبوتية منفصلة عن العنكبوتية الحالية ولكنها تعد امتدادا لها، حيث تكون فيها المعلومات محددة المعنى بما يُمكن الحاسبات والبشر من العمل معا"[1]. كما تعرف العنكبوتية الدلالية بشبكة البيانات Web of data والتي يتم فيها الربط بين الأشياء (مثل الأشخاص والكيانات)[2] في مقابل العنكبوتية الحالية التي تعد شبكة من الوثائق Web of documents. والهدف الأساسي منها هو تمكين الحاسبات من القيام بوظائف متقدمة وتطوير النظم التي يمكن أن تدعم التفاعلات الموثوق بها عبر الشبكة[3]. وفي إطار هذه الشبكة يمكن للتطبيقات الذكية أن تكون قادرة على حجز رحلات الطيران، والفنادق وتحديث السجلات الطبية، وتوفير إجابات واحدة ومحددة لأسئلة بعينها دون الحاجة إلى البحث عن المعلومات أو تدقيق النتائج[4]. وهذا يتطلب أن تكون الصفحات و المصادر المتاحة على العنكبوتية الحالية والمفهومة من قبل البشر مفهومة أيضا من قبل الآلات والبرامج حتى يمكن تشارك البيانات وإعادة استخدامها عبر التطبيقات المختلفة، والمؤسسات، والمجتمعات بما يحقق أهداف العنكبوتية الدلالية. فعلى سبيل المثال، عند البحث في محركات البحث عن كلمة

(1) Berners-Lee,T., Hendler, J. and Lassila, O. The Semantic Web. Scientific American. Vol. 284, No. 5 (May 2001), pp. .34-43.As cited in: Brase, Jan. O. Usage of metadata [online]. Elektrotechnik und Informatik der Universt at Hannover, 2005.p28. Available at: https://pdfs.semanticscholar.org/ea90/19ad62ce1ab487881610fe393e9d0c4da6ce.pdf. [Accessed 29/11/2013].

(2) Sikos, Leslie F. Op. Cit.

(3) W3C. Semantic Web [online]. Available At: https://www.w3.org/2001/sw/wiki/Main_Page [Accessed 27/2/2016].

(4) The Semantic Web in Action [online] / Lee Feigenbaum ... [et. al.]. Scientific American. (19 Jan. 2009) Available at: http://www.scientificamerican.com/article/semantic-web-in-actio/ . [Accessed 20/2/2014].

Jaguar فإنها في ظل العنكبوتية الحالية لن تستطع التفرقة ما بين السيارة من ماركة Jaguar والحيوان المفترس Jaguar [(1)]، ومع جعل هذا البيان مفهوما للآلات ستتمكن محركات البحث من توفير نتائج بحث أفضل.

ولخلق هذه الشبكة من البيانات وجعل البيانات مفهومة للآلات يجب أن يتم تنظيم ونشر البيانات باستخدام صيغ أو قوالب Formats تضيف البنية والمعنى لمحتوى الصفحات على العنكبوتية مع ربط البيانات ذات الصلة بعضها ببعض[(2)] وبعبارة أخرى أن يتم تحديد ووصف البيانات الموجودة في الصفحات المتاحة على العنكبوتية في صورة بيانات مهيكلة Structured Data وإعداد روابط فيما بينها.

وهنا يأتي دور ما وراء البيانات حيث أنها تعمل على جعل البيانات المتاحة على الصفحات بيانات مهيكلة قابلة للفهم بواسطة الآلات، وهذا يتفق مع التعريف الذي وضعه بيرنرز لي لما وراء البيانات والذي نص على أنها بيانات مقروءة آليا عن مصادر العنكبوتية وأشياء أخري[(3)]. وبالتالي فإن ما وراء البيانات تلعب دورا أساسيا في هذا الجيل من العنكبوتية . ويمكن القول بأن فكرة ما وراء البيانات البسيطة للمساعدة في العثور على المصادر أعيد صياغتها تحت شعار البيانات المهيكلة لتحسين أداء محركات البحث " structured data for search engine optimization" [(4)]

ولتوفير بيانات مهيكلة على العنكبوتية الدلالية يتم الاعتماد بشكل أساسي على إطار وصف المصدر RDF. والأنطولوجيات Ontologies والتي تعرف أيضا باسم المفردات vocabularies[(5)]. وسيرد في عنصر معايير ما وراء البيانات تفصيلا لهذه المعايير.

---

(1) Sikos, Leslie F. Op. Cit.

(2) Ibid.

(3) Berners-Lee, Tim . Op Cit.

(4) DCMI Metadata Basics [online]. Available at: http://dublincore.org/metadata-basics/. [Accessed 19/10/2013].

(5) Kinsella, Sheila. Augmenting Social Media Items with Metadata using Related Web Content [online]/ Supervisor John Breslin. Digital Enterprise Research Instituite, National University of Ireland, Galway, January 2012 (Ph.D.). p 34. Available at: https://aran.library.nuigalway.ie/xmlui/bitstream/handle/10379/2674/paper.pdf?sequence=1 .[Accessed 17/2/2016].

ووفقا لما سبق يتضح تواجد ما وراء البيانات عبر أجيال العنكبوتية والحاجة الدائمة إليها كوسيلة لتنظيم المصادر ولكن تظل الاختلافات حول كيفية استخدامها وتوظيفها للقيام بهذا الدور.

2-4 ما وراء البيانات ومحركات بحث العنكبوتية:

تعد محركات البحث من أهم الأدوات لتنظيم محتوى العنكبوتية والإفادة منها، فهي بمثابة نقطة الانطلاق الأولى لدى العديد من المستفيدين يؤكد ذلك أن محرك البحث جوجل Google يحتل الموقع رقم 1 على مستوى العالم وفقا لمعدل زيارات المستفيدين[1]. وتعتمد محركات البحث في عملها على برامج تسمى العناكب Spiders أو الزواحف Crawlers، لبحث محتوى العنكبوتية وجمعه وتكشيفه ويتم تغذية هذه العناكب بقائمة أولية من عناوين المواقع المطلوب زيارتها، لتقوم بعد ذلك بتوسعة القائمة بشكل مستمر عن طريق استخلاص وتتبع المواقع الموحدة للمصادر URLs التي تجدها في الصفحات التي جمعتها، كما تقوم العناكب بشكل دوري بإعادة زيارة المواقع حتى تتمكن من جمع المحتويات الجديدة والمعدلة. وعندما يجد برنامج العنكبوت صفحة تشتمل على ما يستحق جمعه فإنه يقوم بإنشاء واختزان ملخص لتلك الصفحة، ثم يقوم بجمع الكلمات التي توجد بها واختزانها في قاعدة بيانات كشفية، وعندما يقوم المستفيد بإجراء بحث في الإنترنت باستخدام أحد محركات البحث فإنه في واقع الأمر يبحث في تلك القاعدة الكشفية وليس في صفحات العنكبوتية نفسها، وما يتم عرضه هو الملخص الذي سبق اختزانه من قبل؛ ولهذا السبب قد تسترجع في بعض الأحيان صفحات لم يعد لها وجود فعلي على أرض الواقع[2].

(1) Alexa Top 500 Global Sites [online] / Alexa. Available at: http://www.alexa.com/topsites . [Accessed 28/2/2016].

(2) كابلن، بريسيلا. مصدر سابق. ص 154- 155.

وتقدم محركات البحث نتائجها في شكل قوائم ترد فيها روابط للصفحات مع مقتطف غني Rich Snippet وهو النص المختصر الذي يظهر أسفل كل رابط صفحة وترتب هذه النتائج وفقا لخوارزميات تختلف من محرك بحث لآخر، فعلى سبيل المثال يعتمد محرك البحث جوجل بشكل أساسي على خوارزمية تعرف باسم رتبة أو ترتيب الصفحة PageRank التي يتم فيها حساب عدد وجودة الروابط links التي تشير إلى صفحة ما لتحديد تقدير تقريبي لأهمية هذه الصفحة، ويعد تحليل الاستشهادات المرجعية Citation Analysis أحد الأفكار التي تأثر بها ترتيب الصفحة PageRank، وبالإضافة إلى ترتيب الصفحة PageRank فإن جوجل أضاف عبر السنوات العديد من العوامل الأخرى السرية لتحديد ترتيب الصفحات في قوائم النتائج وتم تقديرها بما يزيد عن 250 مؤشر أو عامل مختلف يتم تعديلها بصفة مستمرة[1].

وفي بداية عمل محركات البحث اعتمد العديد منها على ما وراء البيانات الواردة في تيجان الميتا Meta Tags حيث اقترح كل من محرك البحث إنفوسيك Infoseek وألتافيستا AltaVista تاج الكلمات المفتاحية Keywords في عام 1995، وزادت شعبيته بسرعة حتى أصبح واحدا من أكثر عناصر ما وراء البيانات شائعة الاستخدام[2]، كما روج محرك البحث ألتافيستا Alta Vista لتاج الوصف Description [3]، إلا أن العديد من محركات البحث أصبحت تتجاهل ما وراء البيانات نتيجة للممارسات غير السوية

(1) Google Search [online]. In: Wikipedia: The Free Encyclopedia. Available at: https://en.wikipedia.org/wiki/Google_Search .[Accessed 20/12/2015], and PageRank [online]. In: Wikipedia: The Free Encyclopedia. Available at: https://en.wikipedia.org/wiki/PageRank . [Accessed 20/12/2015].

(2) Statistic (June 4, 1997), META attributes by count, Vancouver Webpages, Available at: http://vancouver-webpages.com/META/bycount.shtml. retrieved June 3, 2007. As cited In: Meta element [online]. In: Wikipedia: The Free Encyclopedia. Available at: https://en.wikipedia.org/wiki/Meta_element . [Accessed 11/9/2015].

(3) Gill, Tony. Op Cit.

Spamming مثل إثقال الصفحة أو تحميلها بكم كبير من الكلمات المفتاحية من أجل رفع معدلات استرجاعها من قبل محركات البحث ولإخفاء هذه الكلمات الأساسية عن المستفيد يتم إدخالها بلون خلفية الصفحة، أو كمحتوى لتيجان الميتا[1].

إلا أن هذا التجاهل لا يمكن القطع به، وذلك للأسباب التالية:

1- سرية الخوارزميات التي تستخدمها محركات البحث، كما أن هذه الخوارزميات تختلف من محرك بحث لآخر هذا فضلا عن أنها تتغير باستمرار. فعلى سبيل المثال ذكر محرك البحث جوجل بأنه يقوم بتحديث قواعد الترتيب الخاصة به Ranking Rules كل 48 ساعة. وفي ظل هذه الظروف، فإن فهم نهائي لدور ما وراء البيانات في تيجان الميتا Meta Tags لمحركات البحث هو أمر مستبعد.[2]

2- وجود مجال يعرف بـ "تحسين محركات البحث Search Engine Optimization (SEO) وهو المجال الذي يتعلق بالممارسات أو الإجراءات التي تهدف لزيادة رؤية المواقع Visibility من قبل محركات البحث وكذلك زيادة حركة المرور أو الزيارات Traffic التي تتلقاها المواقع من صفحة نتائج محركات البحث Search Engine Results Page (SERP) ؛ فكلما ورد الموقع في رتبة مرتفعة في صفحة نتائج محرك البحث، كلما ازداد عدد الزيارات التي يتلقاها من مستخدمي محركات البحث[3] ويعد استخدام ما وراء البيانات ممثلة في تيجان العنوان Title، والوصف Description، والكلمات المفتاحية Keywords أحد الممارسات الأساسية الموصى بها في مجال تحسين محركات البحث SEO.

---

(1) كابلن، بريسيلا. مصدر سابق. ص 158.

(2) Meta element [online]. In: Wikipedia: The Free Encyclopedia. Available at: https://en.wikipedia.org/wiki/Meta_element. [Accessed 11/9/2015].

(3) Giomelakis, Dimitrios,Veglis, Andreas. Employing Search Engine Optimization Techniques in Online News Articles [online]. Studies in Media and Communication. Vol. 3, No. 1 (June 2015). Available at: http://dx.doi. org/10.11114/smc.v3i1.683 . [Accessed 26/11/2015].

3- اختلاف استخدام ما وراء البيانات بين محركات البحث وتضارب التصريحات حولها، فعلى سبيل المثال في الوقت الذي أشار فيه جوجل لفهمه لعدد من تيجان ما وراء البيانات مثل تاج العنوان، والوصف، والكلمات المفتاحية إلا أنه صرح بعدم استخدامها في ترتيب الصفحات Ranking في صفحة نتائج البحث، ولكنه يقوم بالاستعانة بها في طريقة أو شكل عرض المواقع في صفحة نتائج محرك البحث وتوفير المقتطف الغني Rich Snippet [1]. وهذا قد يشير إلى أنه بالفعل يتجاهل ما وراء البيانات في جانب وهو الترتيب الطبقي للمواقع إلا أنه يستخدمها في جانب آخر وهو المقتطفات الغنية التي تأتي مصاحبة لكل نتيجة في صفحة نتائج البحث. بينما أعلن ياهو عن دعمه لتاج الكلمات المفتاحية بالاشتراك مع عدد من العوامل الأخرى من أجل تحسين ترتيب نتائج البحث ثم أعلن في أكتوبر من عام 2009 خلال الدائرة المستديرة لمحركات البحث Search Engine Round Table عن إسقاط هذا التاج من حساباته، إلا أنه بعد ذلك تم التصحيح بأن هذا الأمر غير صحيح ؛ حيث أشار المدير المسؤول عن البحث في ياهو Yahoo! بأن " ما تغير في خوارزميات الترتيب لياهو Yahoo! هو أنه في حين أننا لا نزال نقوم بتكشيف تاج الكلمات المفتاحية، فإن تاج الكلمات المفتاحية يحظى بأقل قدر من الأهمية فيما يتعلق بترتيب النتائج وسيكون له تأثير أقل وبخاصة عند تقديم نفس هذه الكلمات في الوثيقة نفسها أو في أي جزء آخر منها"[2].

4- الاتجاه الحالي نحو البيانات المهيكلة Structured data والتي تتوافر من خلال إضافة ما وراء البيانات. وكذلك ظهور بروتوكول الرسم البياني للفيس بوك وبطاقات تويتر التي تمثل في حد ذاتها نفس استخدام تيجان الميتا في بداية محركات البحث - كما سيتضح في الجزء الخاص بالمعايير- كما أن محركات البحث الآن والخوارزميات

(1) Meta tags - Custom Search Help [online]. Available at: https://support.google.com/customsearch/answer/2595557?hl=en. [Accessed 27/5/2014]

(2) Meta element [online]. Op. Cit.

والإمكانات التي تتوافر لها لم تعد مثل تلك التي كانت عليها مع بداية استخدام ما وراء البيانات مما سيجعلها أقل عرضة لعمليات الخداع.

5- وجود عدد من الدراسات التي تشير إلى أهمية ما وراء البيانات لمحركات البحث وأنها تعمل على تحسين نتائج البحث الخاصة بها، كما أن استخدام ما وراء البيانات قد يفتح المجال أمام محركات البحث لتكشيف العنكبوتية الخفية The Hidden Web أو العنكبوتية العميقة The Deep Web التي يقصد بها الكم الهائل من المحتوى المتاح في قواعد البيانات المتخصصة مثل فهارس المكتبات، وبنوك الأخبار التي لا تتوفر إلا كصفحات مهيأة بشكل ديناميكي استجابة لتساؤلات بحث ما، كما أنها تشتمل على صور ومواد سمعية ومرئية وغيرها من الملفات غير النصية الأخرى. ويعد استخدام ما وراء البيانات توجها آخر لإتاحة الوصول لبعض محتويات العنكبوتية الخفية[1]. هذا فضلا عن أن محركات البحث قادرة على تكشيف النصوص دون غيرها، وهي وإن كان يمكنها التعرف على الرسومات والصور، إلا أنها غير قادرة على تحليلها وتفسيرها ما لم تكن مصحوبة بنصوص أو وسائم labels [2] وبالتالي فإن ما وراء البيانات يمكن أن تساهم في توفير الإتاحة والوصول لهذه المصادر.

3-4 معايير ما وراء البيانات في بيئة العنكبوتية:

تتعدد معايير ما وراء البيانات بشكل كبير، ففي إشارة إلى عدد معايير ما وراء البيانات في مجال واحد مثل المكتبات ومؤسسات المعلومات تبين وجود ما يقرب من 105 معيارا، وهذا العدد المتزايد من معايير ما وراء البيانات يعد أمرا مربكا، كما أن العلاقات المتداخلة بينها يزيد من صعوبة الأمر[3].

(1) كابلن، بريسيلا. مصدر سابق. ص 162.

(2) تيلور، ألين ج.. مصدر سابق. ص 28 - 29.

(3) Riley, Jenn. Seeing Standards: A Visualization of the Metadata Universe [online] / content Jenn Riley ; design Devin Becker. 2009 – 2010. Available at: http://www.dlib.indiana.edu/~jenlrile/metadatamap/seeingstandards.pdf . [Accessed 26/12/2015].

ويمكن تقسيم معايير ما وراء البيانات بطرق متعددة ؛ حيث يمكن تقسيمها وفقا لـ:

1- المجتمع Community: أي المجتمعات أو التخصصات التي تستخدم هذا المعيار (مثل المعايير الخاصة بالمكتبات).

2- المجال Domain: حيث يشير إلى أنواع المواد التي يخدمها المعيار مثل: النصوص كالكتب والمقالات، أو المواد الموسيقية أو المرئية.

3- الغرض Purpose: ويقصد به نوع ما وراء البيانات الذي يخدمه المعيار، مثل معايير ما وراء البيانات الخاصة بالوصفية منها أو بالبنائية أو بما وراء بيانات الحفظ أو الحقوق.

4- الوظيفة Function: يشير إلى الدور الذي يلعبه المعيار في إنشاء وتخزين ما وراء البيانات. بعض الوظائف تشمل مجرد تحديد الكيانات الأساسية التي يجب وصفها، والبعض الآخر من المعايير تحدد حقول بعينها، بينما يقوم البعض الآخر بتوفير إرشادات عن كيفية تسجيل عنصر بيانات ما، بينما يحدد البعض الآخر هيكل أو بنية البيانات لتخزينها[1].

ونظرا لأن الدراسة الحالية تركز على معايير ما وراء البيانات الخاصة بالعنكبوتية فقط (المجتمع)، وبخاصة معايير ما وراء البيانات الوصفية (الغرض) التي تصف صفحات العنكبوتية (المجال)، فسوف يتم تقسيم أبرز هذه المعايير وأكثرها شهرة إلى الفئات التالية وفقا لأساس الوظيفة لتشمل ما يلي:

1- الخطط Schemas.

2- النماذج المفاهيمية Conceptual Models.

3- المفردات Vocabularies.

4- لغات الترميز Markup Languages أو البنية Syntax.

وفيما يلي شرح لكل فئة من هذه الفئات وأشهر المعايير الخاصة بها، مع ملاحظة أن الفوارق بين هذه الفئات ليست قاطعة وإنما قد تتداخل فيما بينها.

(1) Ibid.

1-3-4 خطط ما وراء البيانات Schemas:

هي عبارة عن مجموعة من عناصر ما وراء البيانات مع قواعد استخدامها. وقصد بهذه المجموعة أن تحقق هدفا معينا.[1] ومن أبرزها في بيئة العنكبوتية عناصر ما وراء بيانات لغة ترميز النص الفائق HTML Metadata أو ما اشتهر بتيجان الميتا HTML Meta Tags، ومعيار بؤرة دبلن DC، وبروتوكول الرسم البياني المفتوح Open Graph Protocol، وبطاقات تويتر Twitter Card.

(أ) عناصر ما وراء بيانات لغة ترميز النص الفائق HTML Metadata:

يقصد بها عناصر ما وراء البيانات التي تحددها لغة ترميز النص الفائق HTML والتي ترد في جزء الرأس head من الصفحة، وفيما يلي شرح لهذه العناصر وبخاصة العناصر التي وردت في الإصدارة الخامسة HTML5 الصادرة عن اتحاد العنكبوتية W3C كتوصية Recommendation في 28 أكتوبر 2014. وقد تم التركيز على أبرز العناصر التي يمكن النظر إليها باعتبارها عناصر ما وراء بيانات كما تحددها هذه الدراسة.

عنصر العنوان title element: عنوان الصفحة ولا يجب أن يكون للصفحة أكثر من عنوان واحد[2]. ويسجل هذا العنوان في تاج العنوان <title> كالتالي:

<title>عنوان الصفحة</title>

عنصر الميتا Meta element: يعرف أيضا بتيجان الميتا Meta Tags، وهذه العناصر لا تعرض للمستفيد النهائي وإنما تكون مفهومة للآلات وعادة ما ترد في صورة أزواج الاسم / القيمة name/value [3] ويوفر هذا العنصر ما وراء بيانات عن الصفحة مثل

(1) كابلن، بريسيلا. مصدر سابق. ص 46.

(2) W3C. Document metadata — HTML5 [online].Available at: http://www.w3.org/TR/html5/document-metadata.html#the-meta-elemen.[Accessed 12/10/2015].

(3) W3schools.com. HTML <meta> Tag [online]. Available at: http://www.w3schools.com/tags/tag_meta.asp . [Accessed 12/10/2015].

الوصف والكلمات المفتاحية والتي عادة ما ترد في الخاصية name، كما يستخدم في محاكاة تأثير حقل الرأس في بروتوكول نقل النص الفائق HTTP من خلال استخدام الخاصية http-equiv، وأيضا في الإعلان عن أسلوب ترميز المحارف في الصفحة Character encoding declaration عن طريق الخاصية charset[(1)]. وفيما يلي شرح لخصائص attributes هذا العنصر.

الخاصية name:

يحدد في هذه الخاصية اسم عنصر ما وراء البيانات ولابد أن يصاحبها الخاصية content ليسجل فيها بيان عنصر ما وراء البيانات ومن أبرز القيم لهذه الخاصية ما يلي:

- اسم-التطبيق application-name: يسجل فيه اسم التطبيق إذا ما كانت الصفحة الموصوفة عبارة عن تطبيق ما.
- المؤلف author: اسم مؤلف الصفحة.
- الوصف description: وصف الصفحة.
- المولد generator: اسم البرنامج المستخدم في إنشاء الصفحة.
- الكلمات المفتاحية keywords: الكلمات المفتاحية التي تعبر عن محتوى الصفحة، وترد هذه الكلمات مفصولة بفاصلة comma [(2)].
- الروبوت robots: تستخدم هذه الخاصية في تحديد كيفية تعامل زواحف أو عناكب محركات البحث مع الصفحة[(3)].

---

(1) W3C. Op. Cit.

(2) Ibid

(3) البسيوني، بدوية محمد. تيجان الميتاداتا meta tags ومدى تمثيلها في صفحات الويب: دراسة تطبيقية على مواقع التميز الرقمي العربية على الانترنت. الاتجاهات الحديثة في المكتبات والمعلومات. مج 16، ع32 (يوليو 2009). ص 196.

ويمكن النظر إلى هذه القيم باعتبارها عناصر ما وراء البيانات الأساسية التي تحددها لغة ترميز النص الفائق HTML لوصف صفحة ما. مثال:

<meta name="description" content="وصف الصفحة ">

<meta name="keywords" content="الكلمة المفتاحية الأولى ,الكلمة المفتاحية الثانية">

<meta name="author" content="مؤلف الصفحة">

الخاصية http-equiv:

تحاكي هذه الخاصية حقول الرأس في بروتوكول نقل النص الفائق HTTP ولابد أن تصاحبها الخاصية content التي يسجل فيها المحتوى، ومن أبرز قيم هذه الخاصية هي content-type نوع المحتوى والتي يحدد فيها طبيعة محتوى الصفحة والذي عادة ما يكون text/html[1]. وترد هذه الخاصية كالتالي:

<meta http-equiv="content-type" content="text/html">

الخاصية charset:

تستخدم هذه الخاصية للإعلان عن خطة ترميز الحروف في الصفحة[2]. ويوضحها المثال التالي:

<meta charset="UTF-8">

وتجدر الإشارة إلى أنه يمكن النظر إلى الخاصية name بأنها تحتوى على عناصر لما وراء البيانات الوصفية كما تعرفها هذه الدراسة، بينما تحتوي كل من الخاصية http-equiv، والخاصية charset على ما وراء البيانات الفنية.

---

(1) W3C. Op. Cit/

(2) W3schools.com. Op. Cit.

(ب) بؤرة دبلن Dublin Core:

تشير دبلن Dublin إلى مدينة دبلن في ولاية أوهايو بالولايات المتحدة الأمريكية حيث تم ابتكار هذه الخطة عام 1995 في ورشة عمل ما وراء البيانات التي استضافها مركز تحسيب المكتبات على الخط المباشر (Online Computer Library Center (OCLC، والمركز الوطني لتطبيقات الحوسبة الفائقة National Center for Supercomputing Applications (NCSA). بينما تشير كلمة بؤرة Core إلى عناصرها باعتبارها عامة وواسعة النطاق ويمكن استخدامها في وصف مدى واسع من المصادر. وتم إنشاء هذه الخطة ومتابعتها من قبل مجموعة من المتخصصين في تخصصات متعددة تشمل المكتبات، وعلوم الحاسب الآلي، وتكويد النصوص، والمتاحف[1]. وتمت الموافقة على هذه الخطة أو المعيار كمواصفة أمريكية عام 2001 (ANSI/NISO standard z39.82 – 2001)، كما تم اعتماده كمواصفة دولية صادرة عن الأيزو ISO 15836/2003. وقامت الهيئة المصرية العامة للمواصفات والجودة بتعريب هذا المعيار تحت عنوان: معلومات وتوثيق – مجموعة عناصر بيانات دبلن كور للميتاداتا (البيانات الواصفة)، وصدرت هذه الترجمة عام 2006 (م ق م 2006/5146)[2]، وأعيد تعديلها وإصدارها عام 2014 تحت رقم (5146 / 2014).[3]

وتعد بؤرة دبلن خطة ما وراء بيانات عامة قصد بها أن تكون قادرة على وصف أي

(1) Dublin Core Metadata Element Set, Version 1.1 [online]. 14-06-2012. Available at: http://www.dublincore.org/documents/dces / [Accessed 21/3/2016], .and Dublin Core [online]. In: Wikipedia: The Free Encyclopedia. Available at: https://en.wikipedia.org/wiki/Dublin_Core. [Accessed 20/6/2014].

(2) عبد الهادي، محمد فتحي. مصدر سابق. ص 110.

(3) الهيئة المصرية العامة للمواصفات والجودة. 5146 / 2014: معلومات وتوثيق - مجموعة عناصر بيانات دبلن كور للميتاداتا (البيانات الوصفية)، 2014. متاح على: http://www.eos.org.eg/ar/standard/4125 . [تاريخ الزيارة 2016/5/27].

نوع من مصادر المعلومات، ويتم متابعتها من قبل مبادرة بؤرة دبلن لما وراء البيانات Dublin Core Metadata Initiative (DCMI) [1] وهي الآن منظمة غير ربحية موجودة في المكتبة الوطنية بسنغافورا[2]. وتتكون هذه الخطة في مستواها البسيط من 15 عنصرا تعرف باسم مجموعة عناصر بؤرة دبلن لما وراء البيانات Dublin Core Metadata Element Set وتشمل:

العنوان Title، المنشئ Creator، الموضوع Subject، الوصف Description، الناشر Publisher، المساهم Contributor، التاريخ Date، النوع Type، الشكل Format، المحدد Identifier، المصدر Source، اللغة Language، العلاقة Relation، التغطية Coverage، الحقوق Rights، وجميع هذه العناصر اختيارية وقابلة للتكرار، ويخصص لها المحدد الموحد للمصدر http://purl.org/dc/elements/1.1/ والذي يستخدم كفضاء أسماء namespace[3].

وتعد هذه العناصر الخمسة عشر جزءا من مجموعة ضخمة من العناصر التي تعرف باسم مفردات بؤرة دبلن لما وراء البيانات DCMI Metadata Terms والتي تشمل فئات المصادر، وخطط ترميز المفردات Vocabulary Encoding Schemes، وخطط ترميز البنية Syntax Encoding Schemes. [4]

وفي فترة لاحقة لإنشاء العناصر الخمسة عشر الأساسية تم وضع ما عرف بمحددات بؤرة دبلن Dublin Core Qualifiers إلا أن هذه المحددات حلت محلها مفردات بؤرة دبلنDCMI Metadata Terms في عام 2012[5].

---

(1) Jisc Digital Media | Guide [online]: Putting Things in Order: a Directory of Metadata Schemas and Related Standards. Available at: http://www.jiscdigitalmedia.ac.uk/guide/putting-things-in-order-links-to-metadata-schemas-and-related-standards.[Accessed 18/1/2015].

(2) User Guide - DCMI_MediaWiki [online].Available at: http://wiki.dublincore.org/index.php/User_Guide . [Accessed 10/10/2013].

(3) Dublin Core Metadata Element Set, Version 1.1[online]. Op. Cit.

(4) Ibid.

(5) Dublin Core Qualifiers [online]. 11-07-2000. Available at: http://dublincore.org/documents/2000/07/11/dcmes-qualifiers/ . [Accessed 22/10/2013].

ويوضح الجدول التالي رقم (1/1) هذه المفردات والتي يخصص لها المحدد الموحد للمصدر http://purl.org/dc/terms/ [(1)].

وهناك العديد من الطرق لترميز بؤرة دبلن والتي تشمل لغة ترميز النص الفائق HTML، ولغة الترميز الموسعة XML، وإطار وصف المصدر RDF[(2)]. وفي إطار العنكبوتية الدلالية تعد مفردات بؤرة دبلن من المفردات (الأنطولوجيات) الأساسية المستخدمة في إطار البيانات المترابطة Linked data [(3)] كما سيرد ذكره لاحقا.

(ج) بروتوكول الرسم البياني المفتوح (og) Open Graph

هذا البروتوكول تم الإعلان عنه من قبل الفيس بوك Facebook في إبريل 2010 [(4)]. ويتيح لأي صفحة على العنكبوتية بأن تصبح كيانا object في الرسم الاجتماعي Social graph للفيس بوك - الذي يقصد به الخريطة العالمية لجميع الأفراد والعلاقات بينهم حيث تكون الكيانات بمثابة نقاط Nodes والعلاقات تمثل خطوط تسمى Edges [(5)] - وذلك عن طريق إضافة عناصر ما وراء بيانات يحددها هذا البروتوكول لصفحات العنكبوتية في جزء الرأس head.وهذا يُمكن الفيس بوك من التعرف على محتوى هذه الصفحات وبالتالي يكون قادرا على تقديم خدمات متطورة، فعلى سبيل المثال عند الضغط على الزر "إعجاب like" لأحد الصفحات فإنه في حالة عدم توافر عناصر ما وراء البيانات لبروتوكول الرسم البياني

(1) DCMI Metadata Terms [online]. 14-06-2012. Available at: http://dublincore.org/documents/2012/06/14/dcmi-terms/ . [Accessed 20/9/2014]..

(2) DCMI Specifications [online]. Available at: http://dublincore.org/specifications/. [Accessed 19/10/2013]

(3) User Guide - DCMI_MediaWiki [online]. Op. Cit

(4) Yu, Liyang. A Developer's Guide to the Semantic Web [online]. 2nd ed. Berlin: Springer, 2014.p. 520. Available at: http://link.springer.com/content/pdf/10.1007%2F978-3-662-43796-4.pdf . [Accessed 24/3/2016].

(5) The Open Graph protocol [online]. Op. Cit., and Social graph [online]. Op.Cit.

الجدول رقم (1-1) مفردات بؤرة دبلن DCMI Terms

relation العلاقة	Is Part Of جزء من	date التاريخ	abstract الملخص
replaces يحل محل	Is Referenced By استشهد به من قِبل	date Accepted تاريخ القبول	access Rights حقوق الإتاحة
requires يتطلب	Is Replaced By استبدل بـ	Date Copyrighted تاريخ حق النشر	Accrual Method طريقة أو أسلوب الإضافة
rights الحقوق	isRequiredBy مطلوب لـ	Date Submitted تاريخ الإيداع	Accrual Periodicity تتابع الإضافة
Rights Holder صاحب الحقوق	issued صدر	description الوصف	Accrual Policy سياسة الإضافة
Source المصدر	Is Version Of إصدارة من	Education Level مستوى التعليم	alternative العنوان البديل
spatial مكاني	language اللغة	extent المدى	audience الجمهور
subject الموضوع	license الترخيص	format الشكل	available متاح
Table Of Contents قائمة المحتويات	mediator الوسيط	Has Format له شكل	Bibliographic Citation الاستشهاد المرجعي
temporal زمنى	medium الوسيط المادي	Has Part له جزء	Conforms To يتوافق مع
title العنوان	modified تم تعديله	Has Version له إصدارة أو نسخة	contributor المساهم
type النوع	provenance الأصل	identifier المحدد	coverage التغطية
valid تاريخ الصلاحية	publisher الناشر	Instructional Method أسلوب التدريس	created أُنشئ
	references يستشهد بـ	Is Format Of يعد شكلا لـ	creator المنشئ

المفتوح في هذه الصفحة، فإن الفيس بوك يراها مجرد رابط فقط link، بينما في حالة توافر هذه العناصر يكون قادرا على تحديد نوعها، والوصف الخاص بها والأفراد الذين قاموا بالاهتمام بها وبالتالي يمكنه تقديم توصيات حول مزيد من هذا النوع من الصفحات[1]. كما أن توافر هذه العناصر يحدد للفيس بوك طريقة عرض هذه الصفحات في حالة تشاركها shared [2].

ويحدد هذا البروتوكول عددا من العناصر الأساسية وأخرى اختيارية[3] وجميع هذه العناصر تسبقها البادئة og ويوضحها الجدول التالي رقم (1-2):

ويعتمد هذا المعيار في ترميزه على إطار وصف المصدر في الخصائص RDFa وسيرد شرح ذلك في جزء لاحق.

(د) بطاقات تويتر Twitter Cards:

تويتر هو شبكة اجتماعية تتيح لمستخدميها إرسال وقراءة رسائل في حدود 140 تمثيلة، ويطلق على هذه الرسائل تغريدات tweets. وبطاقات تويتر ظهرت في منتصف عام 2012، وتعمل على توفير إمكانية إضافة عناصر ما وراء بيانات مثل الوصف، والعنوان، والصورة لصفحات العنكبوتية التي يمكن أن يشار إليها في هذه الرسائل بما يجعل تويتر قادر على عرض بطاقة عن هذه الصفحات[4].

(1) Yu, Liyang. Op. Cit. p.524.

(2) Facebook. Open Graph Markup [online]. Available at: https://developers.facebook.com/docs/sharing/webmasters#markup . [Accessed 23/3/2016].

(3) The Open Graph protocol [online]. Op. Cit.

(4) Yu, Liyang . Op. Cit. p. 530.

الجدول رقم (2-1) العناصر الأساسية والاختيارية لبروتوكول الرسم البياني المفتوح (og)

العناصر الأساسية	• العنوان: og: title: عنوان الكيان كما يجب أن يظهر في الرسم البياني للفيس بوك. • النوع: og: type: نوع الكيان مثل video.movie، أو موقع website. • الصورة: og: image: الموقع الموحد للمصدر URL للصورة والتي يجب أن تمثل الكيان في الرسم البياني • الموقع الموحد للمصدر URL: هو محدد المكيان الذي سيستخدم كمحدد فريد ودائم لهذا الكيان.
العناصر الاختيارية	• الصوت og: audio: الموقع الموحد للمصدر URL الخاص بملف الصوت المصاحب للكيان. • الوصف og: description: جملة أو اثنين لوصف الكيان. • المحدد og: determiner: الكلمة التي تظهر قبل عنوان الكيان (مثل الحروف a, an,the) • اللغة: og: locale: اللغة الوارد بها هذه التيجان أو العناصر، في شكل (اللغة – كود البلد) والافتراضي هو en_US • اللغة البديلة og:locale:alternate: مجموعة من اللغات الأخرى المتاح بها الصفحة • اسم الموقع og:site_name: إذا كانت الصفحة (الكيان) جزء من موقع أكبر، فيتم استخدام هذا العنصر لعرض اسم هذا الموقع • الفيديو og:video: الموقع الموحد للمصدر URL لملف الفيديو المصاحب للكيان.



الشكل رقم (1-4) بطاقة تويتر Twitter Card لأحد الصفحات أسفل التغريدة tweet

وهناك أربع أنواع من البطاقات تختلف باختلاف نوع المصدر الذي يتم وصفه وهذه الأنواع هي[1]:

1- بطاقة الملخص Summary Card: وهي النوع الأساسي الذي يمكن استخدامه لوصف العديد من الأنواع مثل رسائل المدونات Blog Posts، المقالات، المنتجات.

2- بطاقة الملخص مع صورة كبيرة Summary Card with Large Image: بطاقة مماثلة لبطاقة الملخص لكن مع صورة أكبر.

3- بطاقة التطبيق App Card: بطاقة لوصف تطبيق خاص بالهواتف المحمولة مع رابط التحميل الخاص به.

4- بطاقة المشغل Player Card: بطاقة لوصف وسائط مثل مقاطع الفيديو أو المواد الصوتية.

(1) Twitter. Twitter Cards [online]. Op.Cit.., and Twitter. Summary Card [online]. Available at: https://dev.twitter.com/cards/types/summary .[Accessed 24/3/ 2016].

ويوضح الجدول التالي رقم (1-3) العناصر الأساسية والاختيارية لبطاقة الملخص ويلاحظ أن هذه العناصر دائما ما تسبقها البادئة twitter:

الجدول رقم (1-3) عناصر ما وراء البيانات لبطاقة تويتر الملخص Summary Card

العناصر الأساسية	نوع البطاقة twitter: card: يجب أن يسجل بها القيمة summary اسم الموقع twitter: site: اسم المستخدم صاحب الصفحة أو الموقع الموصوف على تويتر في صيغة @username العنوان twitter: title:عنوان الصفحة الوصف twitter: description: وصف لمحتوى الصفحة
العناصر الاختيارية	الصورة twitter: image: الموقع الموحد للمصدر URL للصورة التي تعبر عن الصفحة ويجب أن تكون أبعاد هذه الصورة 120×120 بكسل

وتجدر الإشارة إلى أنه إذا كانت هناك صفحة من الصفحات لا يوجد بها عناصر بطاقات تويتر واشتملت فقط على عناصر بروتوكول الرسم البياني المفتوح Og، فإن معالج تويتر يمكنه التعرف على هذه العناصر وفهمها والتعامل معها[1].

2-3-4 النماذج المفاهيمية Conceptual Models

النماذج المفاهيمية هي التي توفر التوجه أو التصور الأعلى لوصف المصادر في مجال ما؛ فهي تعمل على تحديد الكيانات والعلاقات فيما بينها. وأشهرها على الإطلاق إطار وصف المصدر RDF[2].

(1) Yu, Liyang . Op. Cit. p.532.

(2) Riley, Jenn. Op. Cit.

(أ) إطار وصف المصدر (RDF) Resource Description Framework:

يعد هذا الإطار أحد الركائز الأساسية للعنكبوتية الدلالية، وتم إنشاءه بواسطة اتحاد العنكبوتية W3C في عام 1999 كمعيار لتمثيل ما وراء البيانات التي تصف المصادر أو الكيانات بما يجعلها مفهومة آليا [(1)]. وجميع الأشياء التي يستخدم إطار وصف المصدر في وصفها تسمى مصادر resources، والمصدر قد يكون صفحة عنكبوتية كاملة، أو قد يكون جزءا من صفحة، أو قد يكون مجموعة من الصفحات مثل موقع كامل، وقد يكون كيان Object غير متاح عبر الإنترنت مثل كتاب مطبوع، وتحدد المصادر دائما بالمحددات الموحدة للمصادر URI [(2)].

ويعتمد هذا الإطار في وصف المصادر على جمل Statements وتعرف أيضا بالثلاثيات Triples ذلك لأنها تتكون من ثلاثة أجزاء – تتماشى مع بنية الجملة في اللغة الإنجليزية - هي:

أ - الفاعل Subject.

ب- الفعل Predicate .

ج- المفعول به Object.

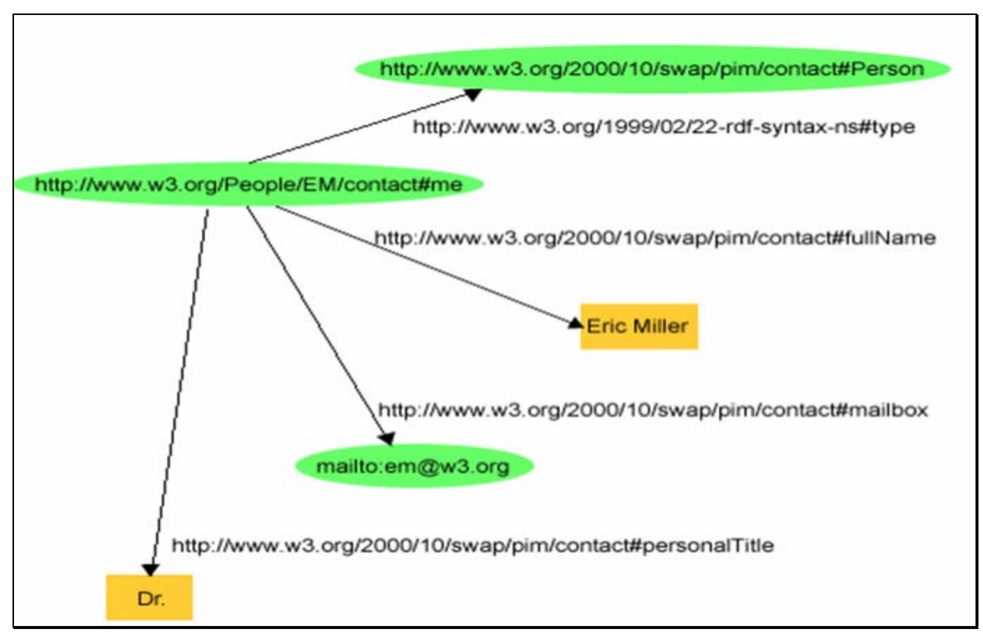
حيث يعبر الفاعل Subject عن المصدر الموصوف، بينما يمثل الفعل Predicate خاصية من خصائص هذا المصدر، ويمثل المفعول به Object قيمة هذه الخاصية وهذه القيمة قد تشكل مصدر آخر قابل لأن يتم وصفه باستخدام جمل إطار وصف المصدر، وهذا يشكل في النهاية شبكة ضخمة من الروابط يتم التعبير عنها في نموذج بياني Graph Model – انظر الشكل رقم (1-5) - ترد فيه كلا من المصادر الموصوفة subject وقيم خصائصها object في شكل نقاط أو عقد nodes، بينما يرد اسم الخاصية predicate أو

---

(1) Yu, Liyang . Op. Cit. p. 24.

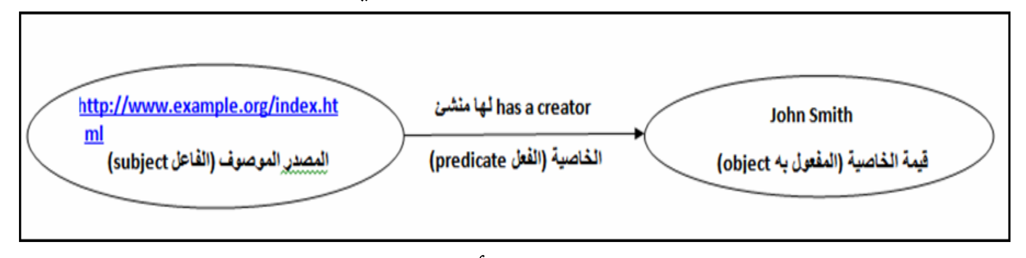
(2) W3C. RDF Primer [online]:W3C Recommendation 10 February 2004. Available at:https://www.w3.org/TR/2004/REC-rdf-primer-20040210/#basicconcepts . [Accessed 28/12/2013].

العلاقة بينهما في شكل أسهم arcs [(1)].



الشكل رقم (1-5) نموذج بياني graph لوصف شخص اسمه إيرك ميلر Eric Miller

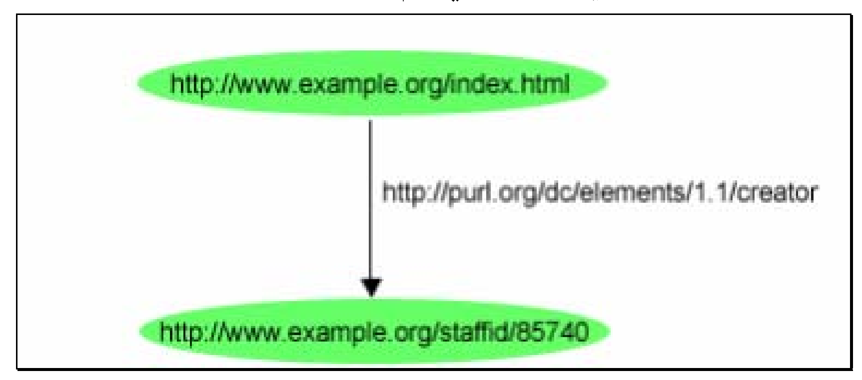
على سبيل المثال: صفحة عنكبوتية أنشأها جون سميث John Smith في هـذه الحالـة تكون الجملة وفقا لإطار وصف المصدر RDF على النحو التالي:



الشكل رقم (1-6) تركيبة الجملة statement (أو الثلاثية triple) في إطار وصف المصدر RDF

(1) Ibid.

وكل عنصر من هذه العناصر الثلاثة يـتم تحديـده بمحـدد موحـد للمصـدر URI حتـى تستطيع الآلات فهم كل عنصر وتحديد معناه بشكل فريد وعادة ما يكون هذا المحدد عبـارة عن الموقع URL، وقد يكون كلا من الفاعل والمفعول عبارة عن قيم حرفية – كما في الشـكل السـابق رقـم (6/1) كانـت قيمـة الخاصـية أو المفعـول بـه John Smith- إلا أن الفعـل (أي الخاصية) يجب أن يكون محدد URI وهذا ما تختص به المفردات أو الأنطولوجيات[1] والتـي عادة ما تحتوي على مجموعة من المفردات يخصص لكل منها محدد URI خاص بهـا. ووفقـا لذلك ترد الجملة السابقة كما في الشكل التالي رقم (1-7):



الشكل رقم (1-7) استخدام مفردات بؤرة دبلن في جملة إطار وصف المصدر RDF

ويلاحظ على المثال السابق أنه تم استخدام خاصية المنشئ creator من مفـردات بـؤرة دبلن، والتي يمكن اختصار كتابة المحدد الخـاص بهـا ليسـجل كالتـالي dc:creator . وفي هـذا إشارة إلى أن بؤرة دبلن يمكن استخدامها كمفردات أو انطولوجيا وهذا ما ستتم الإشـارة إليـه لاحقا. وهذا الإطار يمكن ترميزه باستخدام عدة لغـات أبرزهـا RDF/XML، أو إطـار وصـف المصدر في الخصائص RDFa[2] .

(1) Ibid.

(2) W3C.What's New in RDF 1.1 [online]. Available at: http://www.w3.org/TR/rdf11-new/ ..[Accessed 26/3/2016].

3-3-4 المفردات Vocabularies

يقصد بها قوائم تشمل مفردات أو قيم محددة ومن أمثلتها قائمة النوع لمبادرة بؤرة دبلن لما وراء البيانات DCMI Type. وفي بيئة العنكبوتية الدلالية تقوم المفردات بتحديد المفاهيم والعلاقات المستخدمة في وصف وتمثيل أحد المجالات أو التخصصات، ويتم استخدام مصطلح المفردات بشكل تبادلي مع مصطلح الأنطولوجيات Ontologies حيث أشار اتحاد العنكبوتية W3C إلى أنه لا يوجد فرق واضح بين المصطلحين، إلا أنه في الغالب ما يميل استخدام مصطلح الأنطولوجيا Ontology للتعبير عن المفردات الرسمية والأكثر تعقيدا[1] التي تحتوي على عدد ضخم من المفردات كالمستخدمة في المجال الطبي لوصف الأمراض وأعراضها وعلاجها [2]. وتعد المفردات اللبنة الأساسية لتقنيات الاستدلال وتحديد الدلالات في العنكبوتية الدلالية[3]. وتجدر الإشارة إلى أن هذه المفردات تحدد الفئات والخصائص الواردة بها بمحددات موحدة للمصدر URI حتى يمكن استخدامها في إطار وصف المصدر RDF كما سبق ذكره.

ومن أهم هذه المفردات وأكثرها شهرة سكيما Schema.org، وفواف FOAF، كما أن مفردات معيار بؤرة دبلن DC تستخدم كمفردات في بيئة العنكبوتية الدلالية. وفيما يلي نبذة مختصرة عن كل منها:

(أ) نوع مبادرة بؤرة دبلن لما وراء البيانات DCMI Type:

بصفة عامة تحدد بؤرة دبلن مجموعة من المفردات vocabularies التي توصي باستخدامها مع عدد من عناصرها، والتي من أبرزها نوع مبادرة بؤرة دبلن لما وراء البيانات الموصى باستخدامها مع عنصر النوع Type، وقائمة رؤوس موضوعات مكتبة

(1) W3C. Ontologies - W3C [online]. Available at: https://www.w3.org/standards/semanticweb/ontology .[Accessed 27/6/2014].

(2) W3C. Linked Data Glossary [online]: W3C Working Group Note 27 June 2013. Available at: https://www.w3.org/TR/2013/NOTE-ld-glossary-20130627/ . [Accessed 30/3/2016].

(3) W3C. Ontologies - W3C [online]. Op. Cit.

الكونجرس التي تشكل أحد المفردات الموصى باستخدامها مع عنصر الموضوع Subject، ولكن سيتم التركيز هنا على مفردات النوع التي أعدتها المبادرة، وهذه المفردات تشمل فئات من المفردات التي تعبر عن طبيعة أو نوع المصدر الموصوف، مثل:

مجموعة Collection، مجموعة البيانات Dataset، حدث Event، صورة Image، مصدر تفاعلي Interactive Resource، صورة متحركة Moving Image، كيان مادي Physical Object، خدمة Service، برنامج Software، صوت Sound، صورة ثابتة Still Image، نص Text [1].

(ب) سكيما Schema.org:

هذه المفردات (الأنطولوجيا) صدرت عام 2011[2]، وهي نتاج جهد وتعاون مشترك لمحركات البحث الرئيسية جوجل، وبينج (مايكروسوفت)، وياهو، ومحرك البحث الروسي ياندكس Yandex ويمكن استخدامها في نشر البيانات المهيكلة وتوفير رؤية أفضل للمحتوى أمام محركات البحث وغيرها من التطبيقات وهناك أكثر من 10 مليون موقع يستخدم هذه المفردات[3]. وتتكون هذه المفردات من مجموعة من الأنواع Types (أو الفئات Classes) التي ترد في ترتيب هرمي حيث يحدد نوع المصدر أو الكيان الموصوف بمفردة ما، وكل نوع ترتبط به مجموعة من الخصائص properties، وقد وصل عدد الأنواع أو الفئات إلى 642 نوعا، و992 خاصية[4] ووفقا لذلك يمكن استخدام هذه

(1) DCMI Metadata Terms [online].Op. Cit.

(2) Yu, Liyang. Op. Cit.

(3) Petiya, Sean. Building a Semantic Web of Comics: Publishing Linked Data in HTML/RDFa [online]. (Master Thesis) - Kent State University - College of Communication and Information.2014. Available at: http://rave.ohiolink.edu/etdc/view?acc_num=kent1416791055. [Accessed 29/11/2014], and schema.org. Home - schema.org [online]. Available at: http://schema.org / [Accessed 18/10/2015].

(4) Schema.org . Schemas - schema.org. [online]: Organization of Schemas. Available at: http://schema.org/docs/schemas.html . [Accessed 18/10/2015].

المفردات في وصف مدى واسع من الكيانات أو الفئات مثل الأعمال الإبداعية Creative Work، والأشخاص Person، والمنتجات Product، والأماكن Place وغيرها من الفئات. وتجدر الإشارة إلى أن الأنواع أو الفئات عادة ما يرد الحرف الأول منها بحرف كبير capital، بينما يرد الحرف الأول من الخصائص بحرف صغير small.

على سبيل المثال يمكن وصف المقالات العلمية تحت الفئة أو النوع Scholarly Article باستخدام الخصائص التي ترد معها والتي منها على سبيل المثال لا الحصر: عنوان المقالة name، المؤلف author، تاريخ النشر date Published [1]. وترد هذه المفردات كاملة على موقع سكيما على العنكبوتية schema.org/ .

(ج) صديق الصديق (فواف) Friend of a friend (FOAF):

فواف FOAF هي مفردات أنشئت بواسطة دان بريكلي Dan Brickley وليبي ميلر Libby Miller في منتصف عام 2000، وتستخدم هذه المفردات بصفة أساسية لوصف الأشخاص وأنشطتهم والعلاقات بينهم. وتتكون هذه المفردات مثل أي أنطولوجيا من فئات Classes وخصائص properties، فعلى سبيل المثال فئة الأفراد Person التي تستخدم لوصف شخص ما، ترد تحتها خصائص منها الاسم الأول first name، اسم العائلة surname، يعرف knows، البريد الإلكتروني لهذا الشخص mbox.[2]

(د) بؤرة دبلن Dublin Core (DC):

كما ورد سابقا، فإن بؤرة دبلن بدأت بخمسة عشر عنصرا عرفت بمجموعة عناصر بؤرة دبلن لما وراء البيانات Dublin Core Metadata Element Set، إلا أنه ومع ظهور إطار وصف المصدر RDF أصبحت بؤرة دبلن من أكثر المفردات استخداما مع هذا

---

(1) Schema.org. ScholarlyArticle - schema.org. [online]. Available at: http:// schema.org/ ScholarlyArticle . [Accessed 18/10/2015].

(2) Brickley, Dan and Miller, Libby. FOAF Vocabulary Specification 0.99 [online]. Available at: http://xmlns.com/foaf/spec/. [Accessed 29/3/ 2016].

الإطار[1]، وتمت الاستعانة بالمفردات التي توفرها بؤرة دبلن في وصف مصادر المعلومات سواء المادية أو المتاحة على العنكبوتية [2]. وتعد بؤرة دبلن مفردات أو أنطولوجيا صغيرة نسبيا تنقسم إلى فئتين: مجموعة عناصر بؤرة دبلن، ومفردات بؤرة دبلن DCMI Terms (انظر الجدول رقم 1-1 الوارد سابقا)، حيث تتكون العناصر من الخمسة عشر عنصرا الأساسية بينما تشتمل المفردات على 22 فئة و 55 خاصية [3].

4-3-4 لغات الترميز Markup Language أو البنية Syntax:

هي مجموعة من الرموز أو الأكواد التي تستخدم معا لترميز العناصر بحيث يمكن قراءتها من خلال الحاسبات وتبادلها عبر الشبكات[4]. ونظرا لأن صفحات العنكبوتية عادة ما يتم إعدادها باستخدام لغة ترميز النص الفائق HTML، فكان من الطبيعي استخدام نفس اللغة في ترميز عناصر ما وراء البيانات الواصفة لهذه الصفحات بحيث ترد في داخل الصفحة كجزء من الشفرة code الخاصة بها، ولذلك فإن لغات ترميز ما وراء البيانات في بيئة العنكبوتية تشمل لغة ترميز النص الفائق HTML، وما يرتبط بها في بيئة العنكبوتية الدلالية من إطار وصف المصدر في الخصائص RDFa، والمايكروداتا Microdata.

(أ) لغة ترميز النص الفائق HyberText Markup Language (HTML):

تعد لغة ترميز النص الفائق HTML هي اللغة الأساسية لإنشاء صفحات العنكبوتية،[5]. وتتكون صفحة مكتوبة بلغة ترميز النص الفائق HTML من مجموعة من العناصر elements يتم التعبير عنها من خلال تيجان tags، هذه التيجان قد تصاحبها

---

(1) DCMI Metadata Basics [online]. Op. Cit.

(2) W3C. Linked Data Glossary [online]. Op. Cit.

(3) W3C. Good Ontologies - W3C Wiki [online]. Available at: https://www.w3.org/wiki/Good_Ontologies . [Accessed 29/3/2016].

(4) عبد الهادي، محمد فتحي. مصدر سابق. ص 74، وكابلن، بريسيلا. مصدر سابق. ص 50.

(5) W3C. W3C HTML [online]: The Web's Core Language Available at: https://www.w3.org/html/ . [Accessed 30/3/2016].

خصائص Attributesتوفر معلومات إضافية. مثال:

```
<!DOCTYPE html>
<html>
<head>
<title>Page Title</title>
</head>
<body>
<h1>First Heading</h1>
<p>paragraph.</p>
</body>
</html>
```

ويمكن استخدام لغة ترميز النص الفائق HTML لإضافة عناصر ما وراء البيانات وفقا لخطط ما وراء البيانات المختلفة مثل بؤرة دبلن، أو بطاقات تويتر. حيث يمكن إضافة هذه العناصر باستخدام تاج الميتا <meta> مع الخصائص المحددة لهذا التاج في جزء الرأس <head> ويمكن توضيح ذلك بالأمثلة التالية على عنصر الوصف description:

أ - تاج الوصف وفقا لعناصر ما وراء البيانات كما تحددها لغة ترميز النص الفائق HTML

```
<meta name="description"  content="وصف الصفحة">
```

ب – تاج الوصف وفقا لبؤرة دبلن

```
<meta name="dc.description"  content="وصف الصفحة">
```

ج - تاج الوصف وفقا لبطاقات تويتر

```
<meta name="twitter:description"  content="وصف الصفحة">
```

ويلاحظ على الأمثلة السابقة أنه في حالة استخدام خطة ما وراء بيانات بعينها يجب وضع بادئة prefix تعبر عن هذه الخطة، مثل dc لبؤرة دبلن، أو twitter لبطاقات تويتر.

(ب) إطار وصف المصدر في الخصائص

Resource Description Framework in Attributes (RDFa)

هو أحد توصيات اتحاد العنكبوتية W3C الصادر في 14 أكتوبر 2008 [1]، وفيه تم تحديد مجموعة من الخصائص attributes لتعزيز المعلومات على العنكبوتية بما يجعلها مقروءة آليا [2]. حيث تتيح هذه الخصائص إمكانية إدراج جمل أو ثلاثيات إطار وصف المصدر RDF Triples (انظر الجزء بإطار وصف المصدر في الأجزاء السابقة) في صفحات مكتوبة بلغة ترميز النص الفائق HTML أو أحد مشتقاتها XHTML وهذا يساهم بشكل كبير في جعل المعلومات المتاحة على العنكبوتية مقروءة آليا. وهذا المعيار مدعوم من قبل سكيما Schema.org وتستخدمه محركات البحث مثل جوجل في توفير المقتطف الغني Rich Snippet، ويمكن استخدامه في ترميز معيار وبروتوكول الرسم البياني للفيس بوك [3]، ومعيار بؤرة دبلن[4].

ويتوافر من هذا المعيار إصدارة كاملة تعرف باسم بؤرة إطار وصف المصدر في الخصائص RDFa Core، وهناك إصدارة بسيطة تعرف باسم RDFa lite تشكل الحد الأدنى من الخصائص الكاملة وهي مجموعة بسيطة وسهلة الاستخدام وتغطي نسبة كبيرة من حالات الاستخدام [5]. وفيما يلي يمكن الإشارة إلى عدد من خصائص هذا المعيار على سبيل المثال لا الحصر:

- about: تستخدم لتحديد المصدر الموصوف (تمثل "الفاعل subject " في جملة إطار

---

(1) Yu, Liyang. Op. Cit. p.107.

(2) W3C. RDFa 1.1 Primer - Third Edition [online]: Rich Structured Data Markup for Web Documents . 2015. Available at: https://www.w3.org/TR/rdfa-primer/ . [Accessed 20/6/2015]..

(3) W3C.RDFa. [online]: Linked Data in HTML.. Available at: http://rdfa.info/ .[Accessed 2/4/2016].

(4) DCMI Specifications.[online] Dublincore.org. Op. Cit.

(5) W3C. RDFa Lite 1.1 - Second Edition [online]. 2015. Available at: https://www.w3.org/TR/rdfa-lite/. [Accessed 20/6/2015].

وصف المصدر RDF Statement)، وتجدر الإشارة إلى أن القيمة الافتراضية لهذه الخاصية about تكون محدد صفحة العنكبوتية التي ترد بها هذه البيانات.

- content: تستخدم لوضع قيمة حرفية literal للخاصية (تمثل المفعول به object في جملة إطار وصف المصدر RDF Statement).
- property: تستخدم للتعبير عن العلاقة بين الفاعل subject والمفعول به object (تمثل الفعل predicate (أي الخاصية) في جملة إطار وصف المصدر)
- resource : تستخدم لتحديد إما المصدر الموصوف مثل الخاصية about أو تشكل مفعول به object (أي قيمة الخاصية) يكون عبارة عن مصدر آخر يتم التعبير عنه بمحدد موحد للمصدر URI غير قابل للنقر عليه clickable.
- typeof: تستخدم لتحديد فئة أو نوع المصدر الموصوف.
- vocab: تستخدم لتحديد المفردات vocabulary المستخدمة في الترميز.
- prefix: تستخدم في حالة استخدام أكثر من قائمة مفردات ويتم تسجيل البادئة الخاصة بكل قائمة مع المحدد الخاص بها[1]. ويمكن توضيح ذلك بالمثال التالي:

```
<html>
 <head prefix="og: http://ogp.me/ns#" >
  <meta property="og:title" content="The Trouble with Bob" />
  <meta property="og:type"  content="text" />
  <meta property="og:image" content="http://example.com/alice/bob-ugly.jpg" />
 </head>
 <body vocab="http://purl.org/dc/terms/" prefix="schema: http://schema.org/">
  <div resource="/alice/posts/trouble_with_bob" typeof="schema:BlogPosting">
    <h2 property="title">The trouble with Bob</h2>
```

(1) W3C. RDFa Core 1.1 - Third Edition [online]: Syntax and processing rules for embedding RDF through attributes. 2015. Available at: https://www.w3.org/TR/rdfa-core/ . [Accessed 20/6/2015].

```
<h3 property="creator" resource="#me">Alice</h3>
<div property="schema:articleBody">
  <p>The trouble with Bob is that he takes much better photos than I do:</p>
</div>
</div>
</body>
</html>
```

وتجدر الإشارة إلى أنه يتضح من المثال السابق أنه يمكن استخدام هذه الخصائص لوضع عناصر ما وراء البيانات مع عنصر الميتا <meta>، وفي هذا المثال تم استخدام معيار بروتوكول الرسم البياني للفيس بوك og حيث تم استخدام كلا من الخاصية property لتسجيل اسم الخاصية، والخاصية content لتسجيل قيمة الخاصية، كما تم ترميز عناصر البيانات في عنصر الجسم <body> وفي هذه الحالة يلاحظ أن قيمة الخاصية مثل خاصية العنوان property ="title" تشكل جزء من النص الذي يظهر في الصفحة ويعرض للبشر وفي نفس الوقت يمكن للآلات فهم دلالته ومعناه.

(ج) البيانات الدقيقة أو المايكروداتا Microdata:

هي طريقة مشابهة لإطار وصف المصدر في الخصائص RDFa تهدف إلى إدراج البيانات المهيكلة في صفحات العنكبوتية المكتوبة بلغة ترميز النص الفائق HTML وقام إيان هيكسون Ian Hickson محرر مواصفات الإصدارة الخامسة للغة ترميز النص الفائق HTML5 بإصدار أول عمل يتعلق بالمايكروداتا في عام 2009 [1]، وتتكون المايكروداتا من مجموعة من الخصائص التي تصف الكيانات. والخصائص الرئيسية لهذا الترميز هي:

(1) Ronallo, Jason. HTML5 Microdata and Schema.org. [online].The Code4lib Journal. Issue 16 (03/02/2012). Available at: http://journal.code4lib.org/ articles/6400 . [Accessed 29/5/2014].

- itemscope: وفقا للمايكروداتا فإن الكيانات الموصوفة بها تسمى وحدات items، وللدلالة على أن هناك كيان بعينه سيتم وصفه يجب أن تسجل الخاصية itemscope، ويتم تكرار هذه الخاصية مع كل كيان جديد.
- itemtype: يحدد فيها نوع الكيان الموصوف.
- itemprop: يحدد فيها اسم الخاصية التي تصف الكيان أو المصدر الموصوف[1].

ويمكن توضيح ما سبق بالمثال التالي:

```
<"div itemscope itemtype="http://schema.org/Organization>
<span itemprop="name">code4lib</span>
<div/>
```

في هذا المثال يتم وصف كيان عبارة عن منظمة، وتم تحديد خاصية الاسم لهذه المنظمة وقيمتها هي code4lib، وبالتالي فهذا النوع من الترميز يعمل مثله مثل إطار وصف المصدر في الخصائص RDFa على أن النص المعروض للبشر يتم ترميزه في نفس الوقت ليكون مفهوما للآلات.

ويدعم هذا النوع من الترميز مفردات سكيما Schema.org إلا أنه من الصعب استخدامه عند الرغبة في المزج بين مفردات مختلفة vocabulary [2].

4-4 أدوات ما وراء البيانات:

يقصد بأدوات ما وراء البيانات التطبيقات والبرامج الآلية التي يمكن الاستعانة بها في إنشاء أو تحليل ما وراء البيانات في بيئة العنكبوتية، ووفقا لذلك يمكن تقسيم هذه الأدوات إلى: أدوات إنشاء ما وراء البيانات، وأدوات لتحليل ما وراء البيانات. وفيما يلي نبذة عن كل منهما:

(1) Ronallo, Jason. Op. Cit.

(2) Schema.org . FAQ - schema.org [online]. Available at: http://schema.org/docs/faq.html .[Accessed 18/10/2015].

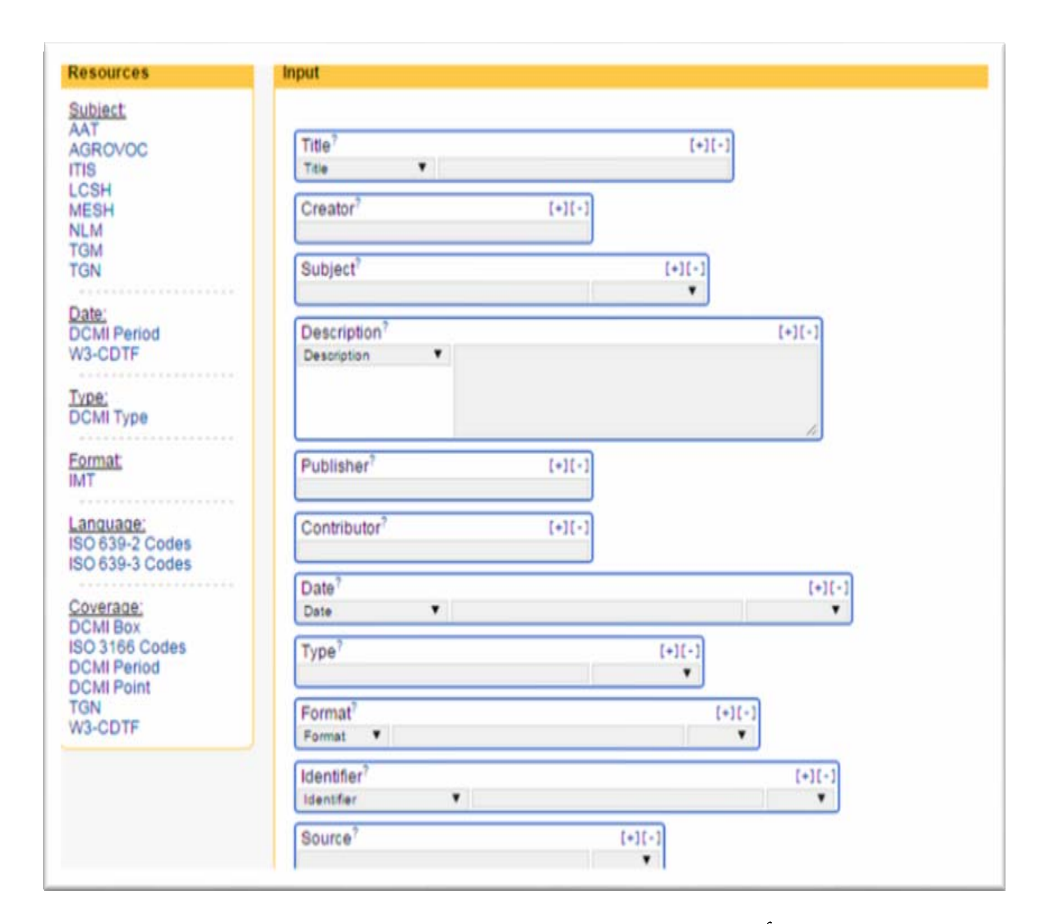
1-4-4 أدوات إنشاء ما وراء البيانات:

عادة ما يكون هذا النوع من الأدوات في شكل استمارة بعناصر ما وراء البيانات، يتم فيها إدخال قيم هذه العناصر، وبناء على ذلك يتم توليد أو إنشاء ترميز بهذه العناصر في شكل قابل لأخذ نسخة منه copy ثم لصقه paste في شفرة المصدر Source code لصفحة العنكبوتية، وتختلف هذه الأدوات فيما بينها من حيث أن بعضها يتاح التعامل معه مباشرة من خلال صفحات على الإنترنت، والبعض الآخر يتطلب أن يتم تحميله على الحاسب الآلي، فيما يشكل بعضها إضافات plugins لبرامج إدارة المحتوى CMS مثل دروبال Drupal والوردبريس WordPress التي تستخدم في إنشاء المواقع.

كما تختلف هذه الأدوات من حيث خطط ما وراء البيانات التي تدعمها، والبنية أو الترميز الذي يتم فيه توليد وإنشاء هذه العناصر؛ فبعضها يوفر ترميز عناصر ما وراء البيانات وفقا لبؤرة دبلن، ومثال ذلك أداة The Advanced Dublin Core Generator [1] (انظر الشكل رقم 1-8).

---

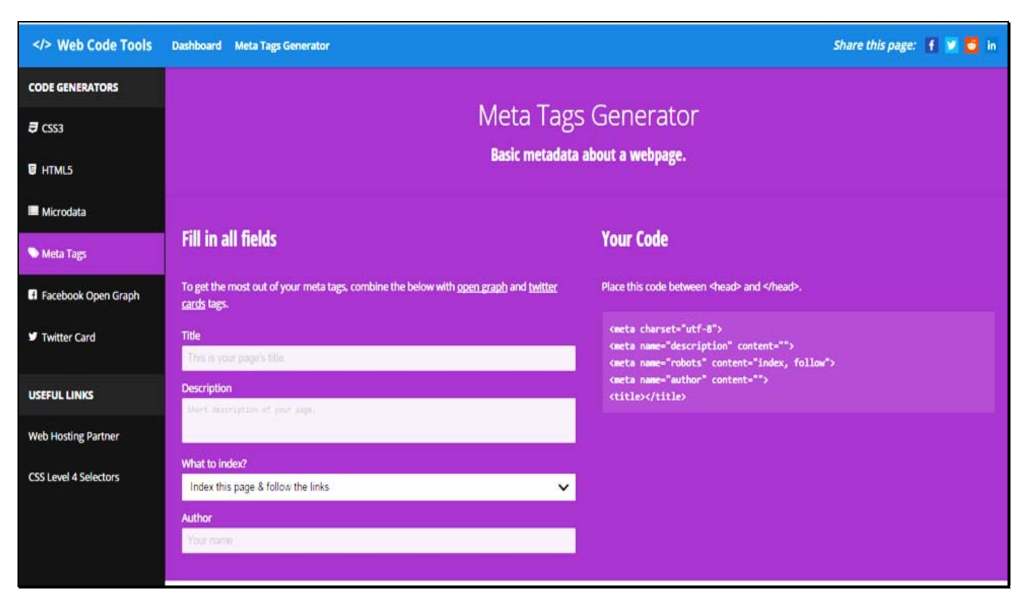
(1) The Advanced Dublin Core Generator! [online]. Available at . http://www.dublincoregenerator.com/generator.html .[Accessed 10/8/2015].



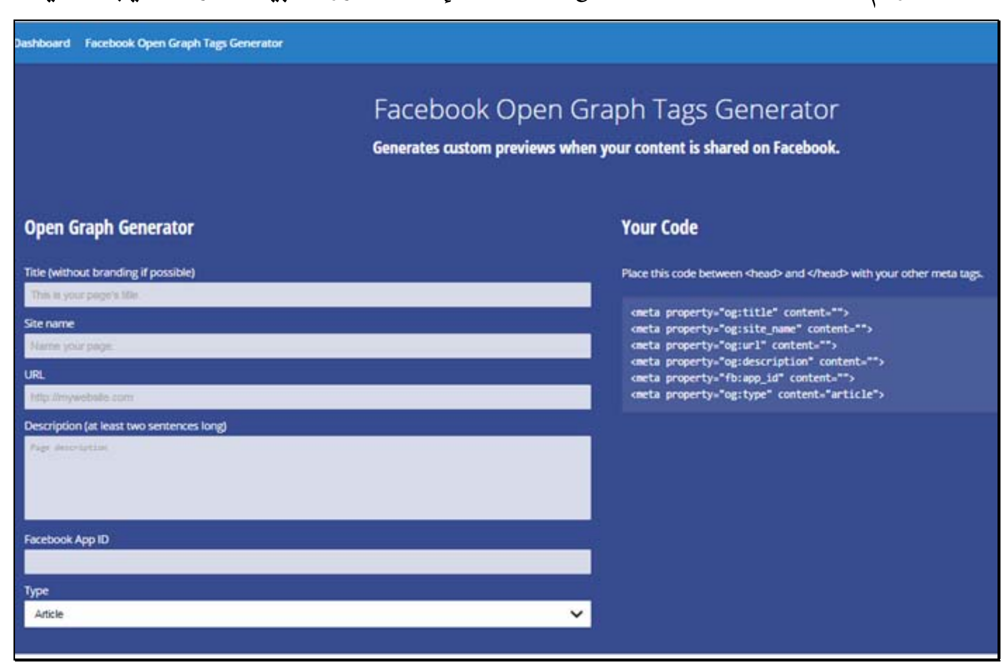
الشكل رقم (1-8) أداة The Advanced Dublin Core Generator لإنشاء ما وراء البيانات وفقا لبؤرة دبلن

بينما يعمل البعض الآخر على توفير عناصر ما وراء البيانات وفقا لتيجان الميتا كما توصي بها لغة ترميز النص الفائق، أو وفقا لمعيار بروتوكول الرسم البياني، أو وفقا لبطاقات تويتر مثال ذلك الأدوات المتاحة على موقع Web Code Tools [1].

(1) WebCodeTools[online]. Available at: http://webcodetools.com/ . .[Accessed 10/8/2015].



الشكل رقم (1-9) أداة Meta Tags Generator لإنشاء ما وراء البيانات وفقا لتيجان الميتا



الشكل رقم (1-10) أداة Facebook Open Graph Generator لإنشاء ما وراء البيانات

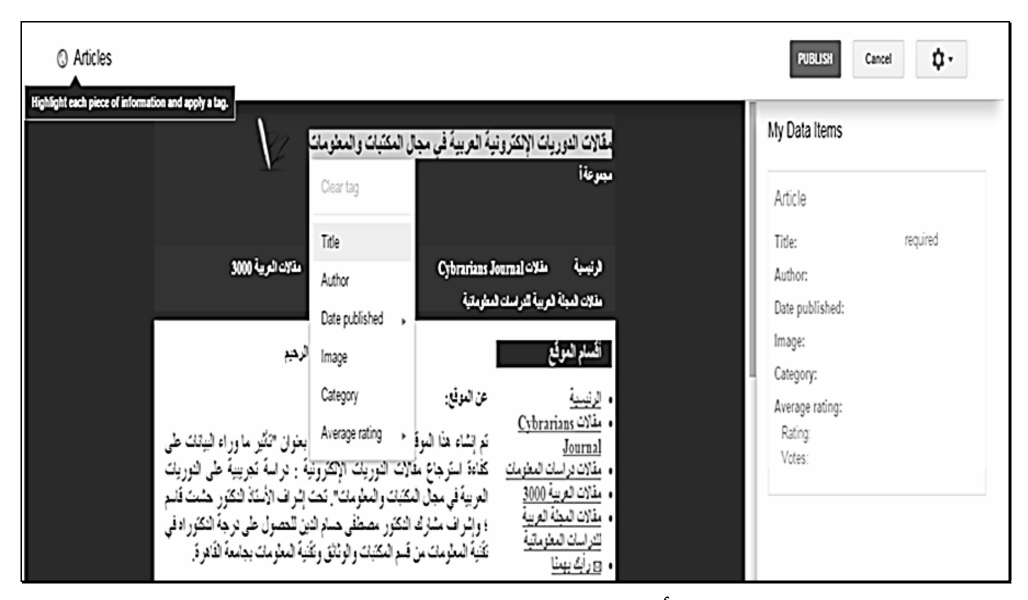
وفقا لبروتوكول الرسم البياني للفيس بوك

الشكل رقم (11-1) أداة Twitter Card Generator لإنشاء ما وراء البيانات وفقا لبطاقات تويتر

وفي إطار اتجاه العنكبوتية الحالية إلى العنكبوتية الدلالية، تجدر الإشارة إلى أحد الأدوات التي يوفرها محرك البحث جوجل والتي يمكن اعتبارها بمثابة أداة لإنشاء ما وراء البيانات .هذه الأداة تعرف باسم "محدد البيانات Data Highlighter " وهي خاصة بجوجل فقط، ولا يمكن أن تفيد منها محركات البحث الأخرى[1]؛ حيث تعمل هذه الأداة على إخبار جوجل بعناصر ما وراء البيانات الموجودة في صفحة ما عن طريق تحديد highlight عناصر مثل العنوان، أو المؤلف في النص الموجود في الصفحة باستخدام الفارة وهذا يساعد جوجل على عرض هذه الصفحة بشكل أكثر جاذبية في نتائج البحث (انظر الشكل رقم 1-12)[2]. وهذه الأداة تتطلب أن يكون لصاحب الموقع حساب account على أدوات جوجل لمصممي المواقع Webmasters tools.

(1) Starr, Barbara. Google's Hunger For Structured Markup [online]. Search Engine Land, 2013. Available at: http://searchengineland.com/googles-hunger-for-structured-markup-160240 . [Accessed 5/4/2016].

(2) Google. About Data Highlighter - Search Console Help [online]. Available at: https://support.google.com/webmasters/answer/2692911?hl=en . [Accessed 10/4/2016].



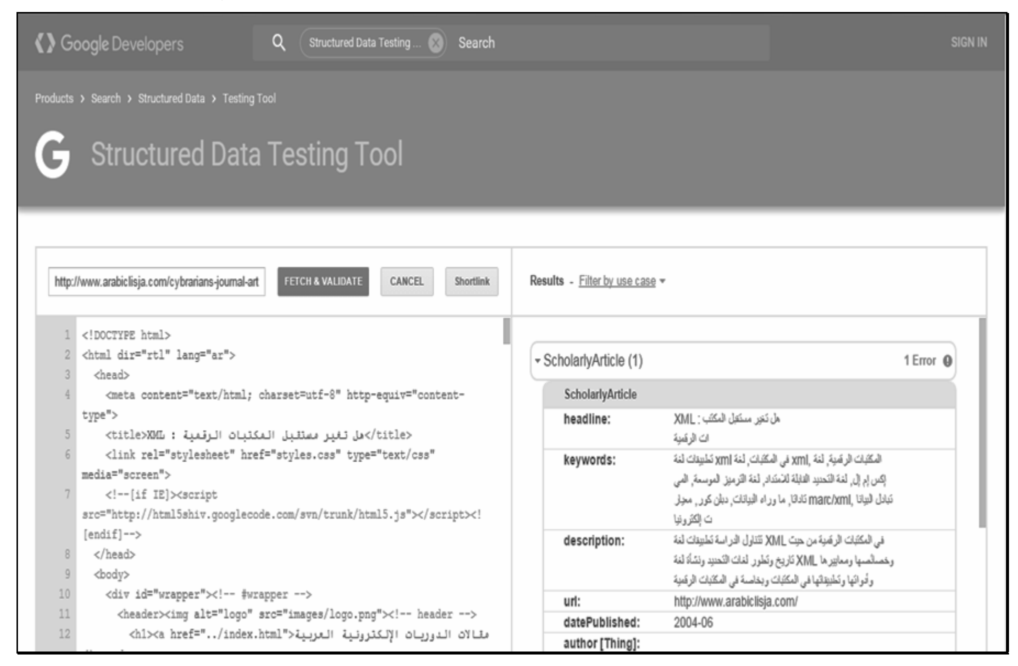
الشكل رقم (1-12) أداة محدد البيانات Data Highlighter لجوجل

2-4-4 أدوات تحليل ما وراء البيانات:

تقوم هذه الأدوات بتحليل عناصر ما وراء البيانـات التـي تتـوافر في صـفحة مـا بهـدف اختبار صحة هذه العناصر وتحديد مدى دقتها. وعادة ما يتم ذلك عن طريق تسجيل الموقع الموحد للمصدر URL للصفحة المراد تحليل ما وراء البيانات الخاصة بها، والبعض الآخـر مـن هذه الأدوات يتيح بجانب ذلك إمكانية الاختبار عن طريـق إضـافة شـفرة المصـدر Source Code لصفحة ما ويمكن استخدام هذه الإمكانية عند الرغبة في اختبـار ترميـز الصـفحة قبـل وضعها عـلى موقـع عـلى الإنترنـت. ومـن أمثلـة هـذه الأدوات وأشـهرها برنـامج ميتـا شـيكر MetaChecker [1]

(1) Meta Tag Analysis Tool by MetaChecker.net [online]. Available at: http://www.metachecker.net/ . [Accessed 15/11/2015].

كما يوفر عدد من محركات البحث مثل جوجل وياندكس Yandex أدوات لتحليل وعرض ما وراء البيانات المعتمدة على إطار وصف المصدر RDF،حيث يوفر جوجل أداة اختبار البيانات المهيكلة Structured Data Testing Tool [(1)] المتاحة لاختبار صحة هذه البيانات ويمكن ذلك عن طريق إدخال الموقع الموحد للمصدر الخاص بالصحفة URL، أو عن طريق إدخال شفرة المصدر الخاصة بهذه الصفحة كما يتضح في الشكل رقم (1-13).



الشكل رقم (1-13) أداة جوجل Structured Data Testing Tool لاختبار صحة ما وراء البيانات

(1) Google Developers. Structured Data Testing Tool [online]. Available at: https://developers.google.com/structured-data/testing-tool/. [Accessed 7/4/2016].

كما يتيح محرك البحث الروسي ياندكس Yandex أداة مشابهة لأداة جوجل تعرف باسم أداة التحقق من صحة البيانات المهيكلة Structured Data Validator [1] كما تظهر في الشكل رقم (1-14).

الشكل رقم (1-14) أداة ياندكس للتحقق من صحة البيانات المهيكلة

Structured Data Validator

(1) Yandex webmaster. Structured data validator [online]. Available at: https://webmaster.yandex.com/microtest.xml . .[Accessed 7/4/2016].

خلاصة الفصل:

حاول هذا الفصل استعراض أبرز الجوانب النظرية حول موضوع ما وراء البيانات مع التركيز على ما وراء البيانات في بيئة العنكبوتية بصفة خاصة، ويمكن تلخيص أبرز هذه الجوانب في النقاط التالية:

1- يرجع تاريخ مصطلح ما وراء البيانات إلى أواخر الستينات من القرن العشرين الميلادي، وينتمي في الأصل إلى علم الحاسب الآلي.

2- يتكون مصطلح ما وراء البيانات Metadata من عنصرين هما **meta** من اليونانية وتعني "ما وراء"، أو "ما بعد"، أو "عن"، و**data** من اللاتينية وتعني البيانات، وهذا المصطلح في أوسع معانيه يعني "البيانات عن البيانات".

3- تتعدد تعريفات ما وراء البيانات نظرا لاستخدام هذا المصطلح من قبل عدد كبير من التخصصات والبيئات المختلفة.

4- تضع هذه الدراسة تعريفا لما وراء البيانات ينص على أنها: بيانات مهيكلة عن أي كيان (سواء كان مصدر معلومات – رقمي أو غير رقمي -، شخص، هيئة، خدمة... إلخ)، وذلك لتحقيق أهداف متنوعة مثل التعريف والاسترجاع والاستخدام والإدارة.

5- تقسم ما وراء البيانات بصفة أساسية إلى ثلاث فئات واسعة هي: ما وراء البيانات الوصفية Descriptive Metadata، وما وراء البيانات البنائية Structural Metadata، وما وراء البيانات الإدارية Administrative Metadata.

6- تضع هذه الدراسة تقسيما لأنواع ما وراء البيانات وفقا للبيئة التي تعمل فيها وبما يتناسب مع طبيعة هذه الدراسة وقد شمل هذا التقسيم الأنواع التالية: ما وراء بيانات العنكبوتية، وما وراء البيانات الببليوجرافية، وما وراء بيانات نظم إدارة قواعد البيانات، وما وراء بيانات البرامج.

7- لما وراء البيانات دورا بارزا ظهر عبر أجيال العنكبوتية، من الجيل الأول Web 1.0 حيث تم التفكير فيها كوسيلة للعثور على واكتشاف المصادر وبخاصة مع زيادة أعداد

المصادر وانتشار استخدام الإنترنت، وفي الجيل الثاني Web 2.0 ظهرت ما وراء البيانات في شكل التوسيم الاجتماعي، كما استعانت بها الشبكات الاجتماعية مثل الفيس بوك وتويتر كوسيلة لتحديد ذاتية المصادر التي يتم تشاركها عليها وكأحد الطرق التي تحدد طريقة عرض هذه المصادر، وفي الجيل الثالث Web 3.0 تعود ما وراء البيانات تحت شعار ما يطلق عليه الآن البيانات المهيكلة Structured data لجعل البيانات المتاحة على العنكبوتية مفهومة لكل من البشر والآلات.

8- تشمل معايير ما وراء البيانات في بيئة العنكبوتية كل من: خطط ما وراء البيانات Schemas مثل بؤرة دبلن، وتيجان الميتا، والنماذج المفاهيمية مثل إطار وصف المصدر RDF، والمفردات Vocabularies مثل سكيما Schema.org، ولغات الترميز أو البنية مثل لغة ترميز النص الفائق HTML، وما يرتبط بها من إطار وصف المصدر في الخصائص RDFa.

9- تتنوع أدوات ما وراء البيانات على العنكبوتية لتشمل أدوات إنشاء مثل أداة The Advanced Dublin Core Generator، وأدوات تحليل مثل برنامج ميتاشيكر Metachecker.

الفصل الثاني

تحليل ما وراء البيانات في مقالات الدوريات الإلكترونية العربية في مجال المكتبات والمعلومات

ويشتمل هذا الفصل على النقاط التالية:

- ✍ عينة الدراسة.
- ✍ تحليل عناصر ما وراء البيانات في مقالات الدوريات الإلكترونية العربية في مجال المكتبات والمعلومات.
- ✍ نظرة استكشافية لعناصر ما وراء البيانات في عدد من مقالات الدوريات الإلكترونية الأجنبية.

الفصل الثاني

تحليل ما وراء البيانات في مقالات الدوريات الإلكترونية العربية في مجال المكتبات والمعلومات

تمهيد:

يحاول هذا الفصل التعرف على واقع ما وراء البيانات في مقالات الدوريات الإلكترونية العربية في مجال المكتبات والمعلومات المنشورة حتى يناير من عام 2015 – تاريخ الانتهاء من تجميع المقالات للتحليل وتصميم التجربة -، وذلك في خمس دوريات هي: "Cybrarians Journal"، و"العربية 3000"، و"دراسات المعلومات"، و"المجلة العربية للدراسات المعلوماتية"، و"مجلة المكتبات والمعلومات". وتمت الاستعانة بقائمة مراجعة لتحليل عناصر ما وراء البيانات الواردة في مقالات هذه الدوريات يدويا بالاطلاع على شفرة المصدر source code للصفحات التي تتاح عليها هذه المقالات، بالإضافة إلى الاستعانة ببرنامج تحليل ما وراء البيانات ميتاشيكر Metachecker.net. فضلا عن استكشاف عناصر ما وراء البيانات في عينة من مقالات ثلاث دوريات إلكترونية أجنبية هي كل من: "D –Lib Magazine"، و"Ariadne"، و"First Monday".

1 عينة الدراسة:

1-1 الدوريات الإلكترونية العربية في مجال المكتبات والمعلومات محل الدراسة:

الجدول رقم (2-1) الدوريات الإلكترونية العربية في مجال المكتبات والمعلومات محل الدراسة

الدورية	العربية 3000	مجلة المكتبات والمعلومات	Cybrarians Journal	دراسات المعلومات	المجلة العربية للدراسات المعلوماتية
تاريخ صدور العدد الأول	يناير 2000	إبريل 2002	يونيو 2004	يونيو 2008	يوليو 2012
هيئة الإصدار	نادي الإحياء العربي – سوريا	مخبر البحث "تكنولوجيا المعلومات ودورها في التنمية الوطنية"- الجزائر	البوابة العربية للمكتبات والمعلومات – مصر	جمعية المكتبات السعودية – السعودية	المركز العربي للدراسات والبحوث بالتعاون مع معهد الملك سلمان للدراسات والخدمات الاستشارية / جامعة المجمعة - السعودية
تتابع الصدور	فصلية	نصف سنوية	فصلية	مرتين في السنة	مرتين في السنة
إتاحة النص الكامل للمقالات	مجانية	مجانية	مجانية	مجانية بشرط التسجيل في موقع الدورية	باشتراك

الدورية	العربية 3000	مجلة المكتبات والمعلومات	Cybrarians Journal	دراسات المعلومات	المجلة العربية للدراسات المعلوماتية
الأشكال المتاحة	إلكتروني وورقي	إلكتروني وورقي	إلكتروني فقط	إلكتروني وورقي	إلكتروني وورقي
صيغ عرض المقالات	الملخص والنص الكامل بلغة ترميز النص الفائق HTML	الملخص بلغة ترميز النص الفائق HTML، والنص الكامل بصيغة PDF	الملخص والنص الكامل بلغة ترميز النص الفائق HTML	الملخص بلغة ترميز النص الفائق HTML، والنص الكامل بصيغة PDF	الملخص بلغة ترميز النص الفائق HTML، والنص الكامل بصيغة PDF
إصدارة لغة ترميز موقع الدورية	XHTML 1.0 Transitional	XHTML 1.0 Strict	HTML 4.01 Transitional	HTML 5	HTML 5
الموقع الإلكتروني	http://alarabiclub.org/index.php?p_id=212	http://www.webreview.dz-/?مجلة-المكتبات-و-المعلومات	http://journal.cybrarians.info/	http://journals.psiscs.com.sa/index.php/ijs/index (1)	http://journals.psiscs.com.sa/index.php/asj/index (2)

(1) تغير هذا الموقع عن الوقت الذي تمت فيه الدراسة إلى: http://journals.ksiscs.com.sa/index.php/ijs/index . تاريخ الاطلاع [2017/5/8].

(2) تغير هذا الموقع عن الوقت الذي تمت فيه الدراسة إلى: http://journals.ksiscs.com.sa/index.php/asj/ . تاريخ الاطلاع [2017/5/8].

يتضح من الجدول السابق رقم (2-1) ما يلي:

- بلغ عدد الدوريات الإلكترونية العربية في مجال المكتبات والمعلومات التي خضعت للتحليل خمس دوريات إلكترونية، كان أقدمها من حيث صدور أول عدد هي العربية 3000 والتي صدر عددها الأول في عام 2000، وأحدثها المجلة العربية للدراسات المعلوماتية بصدور عددها الأول في عام 2012.
- جميع الدوريات الإلكترونية محل الدراسة تتاح في كل من الشكلين الورقي والإلكتروني فيما عدا دورية Cybrarians Journal والتي تتاح في الشكل الإلكتروني فقط.
- تتيح كل من الدوريات: العربية 3000، ومجلة المكتبات والمعلومات، و Cybrarians Journal النص الكامل لمقالاتها مجانا، غير أن دورية دراسات المعلومات تتيح النص الكامل مجانا بشرط التسجيل في موقع الدورية. أما المجلة العربية للدراسات المعلوماتية فتتيح الملخص فقط مجانا أما النص الكامل للمقالات فيتاح باشتراك.
- تنوعت هذه الدوريات من حيث صيغ إتاحة المقالات على الشبكة العنكبوتية ما بين دوريات تتيح كلا من النص الكامل والملخص بلغة ترميز النص الفائق HTML وذلك كما في الدوريتين العربية 3000، وCybrarians Journ. ودوريات تتيح الملخص وحده بلغة ترميز النص الفائق HTML بينما يتاح النص الكامل للمقالة بصيغة PDF، وذلك في الدوريات مجلة المكتبات والمعلومات، ودراسات المعلومات، والمجلة العربية للدراسات المعلوماتية، وهو ما يمثل العدد الأكبر من الدوريات عينة الدراسة.
- استخدمت كل من دورية دراسات المعلومات والمجلة العربية للدراسات المعلوماتية أحدث إصدارات لغة ترميز النص الفائق وهي الإصدارة الخامسة HTML 5 في ترميز موقعها، فيما استخدمت Cybrarians Journal الإصدارة الرابعة من هذه اللغة HTML 4.01 Transitional بينما استخدمت كل من دورية العربية 3000، ومجلة المكتبات والمعلومات أحد مشتقات هذه اللغة وهي لغة ترميز النص الفائق الموسعة XHTML.
- تجدر الإشارة إلى أن كل من دراسات المعلومات والمجلة العربية للدراسات المعلوماتية

تتم إتاحتهما على موقع واحد هو موقع معهد الملك سلمان للدراسات والخدمات الاستشارية التابع لجامعة "المجمعة"، وذلك بجانب عدد آخر من الدوريات المتخصصة في مجالات أخرى مثل المجلة العربية للدراسات التربوية والاجتماعية، والمجلة العربية للدراسات الإدارية والاقتصادية، والمجلة العربية للدراسات الشرعية والقانونية.

2-1 عينة مقالات الدوريات الإلكترونية العربية في مجال المكتبات والمعلومات:

يعرض الجدول رقم (2-2) عينة الدراسة من المقالات المتاحة على مواقع الدوريات الإلكترونية العربية في المجال والتي خضعت لتحليل ما وراء البيانات بها.

الجدول رقم (2-2) عينة الدراسة من مقالات الدوريات الإلكترونية العربية في مجال المكتبات والمعلومات حتى يناير 2015

الدورية	الأعداد المتاحة حتى يناير 2015	إجمالي المقالات	المقالات عينة الدراسة	نصيب كل دورية من المقالات في عينة الدراسة
Cybrarians Journal	36	276	108	%48.65
العربية 3000	17	227	51	%22.97
دراسات المعلومات	14	106	42	%18.92
المجلة العربية للدراسات المعلوماتية	5	25	15	%6.76
مجلة المكتبات والمعلومات	2	19	6	%2.70
الإجمالي	74	653	222	%100

يتضح من الجدول السابق رقم (2-2) أن:

- بلغ إجمالي عدد مقالات الدوريات الإلكترونية محل الدراسة حتى يناير 2015 بعد استبعاد الافتتاحيات (653) مقالة وردت في (74) عددا، وتم تحديد عينة الدراسة بواقع ثلاث (3) مقالات من كل عدد من الأعداد الصادرة عن كل دورية والمتاحة على موقعها، وتم اختيار هذه المقالات عشوائيا؛ حيث تم اختيار أول ثلاث مقالات

في قائمة المحتويات من بعد الافتتاحية. ووفقا لذلك يصبح إجمالي حجم عينة المقالات (222) مقالة. وبالاستعانة ببرنامج Sample Size Calculator [1] تبين أن درجة الثقة في هذه العينة من المقالات تبلغ 95%، وبهامش خطأ 5.35%.

- حظيت دورية Cybrarians Journal بأكبر عدد من المقالات التي تم تحليلها في عينة الدراسة بعدد (108) مقالة وبنسبة بلغت نحو 48.65% من إجمالي المقالات محل الدراسة، تلتها كل من العربية 3000 بعدد (51) مقالة شكلت نسبة حوالي 22.97%، ثم دراسات المعلومات بعدد (42) مقالة أي بنسبة 18.92%، ثم المجلة العربية للدراسات المعلوماتية بـ(15) مقالة بنسبة 6.76%، وأخيرا مجلة المكتبات والمعلومات بست (6) مقالات بما يعادل حوالي 2.70%؛ هذا ويرجع الاختلاف في عدد المقالات التي تم تحليلها من كل دورية إلى عدد الأعداد والمقالات الصادرة عن الدورية والمتاحة على موقعها، فكلما ارتفع عدد المقالات الصادرة عن هذه الدورية، ارتفع عدد مقالات العينة لهذه الدورية.

2 تحليل عناصر ما وراء البيانات في مقالات الدوريات الإلكترونية العربية في مجال المكتبات والمعلومات.

تم تحليل عناصر ما وراء البيانات في مقالات الدوريات الإلكترونية عينة الدراسة يدويا اعتمادا على قائمة مراجعة، وذلك من خلال فتح شفرة المصدر Source Code لكل مقالة، ويوضح الشكل رقم (2-1) شفرة المصدر لأحد مقالات الدراسة.

كما تمت الاستعانة ببرنامج ميتاشيكر Metachecker.net بجانب التحليل اليدوي، ويقوم هذا البرنامج بتحليل عناصر ما وراء البيانات في موقع ما عن طريق تسجيل الموقع الموحد للمصدر URL، كما يظهر في الشكل التالي رقم (2-2).

(1) Raosoft. Sample Size Calculator by Raosoft, Inc [online]. Available at: http://www.raosoft.com/samplesize.html.[Accessed 24/5/2014].

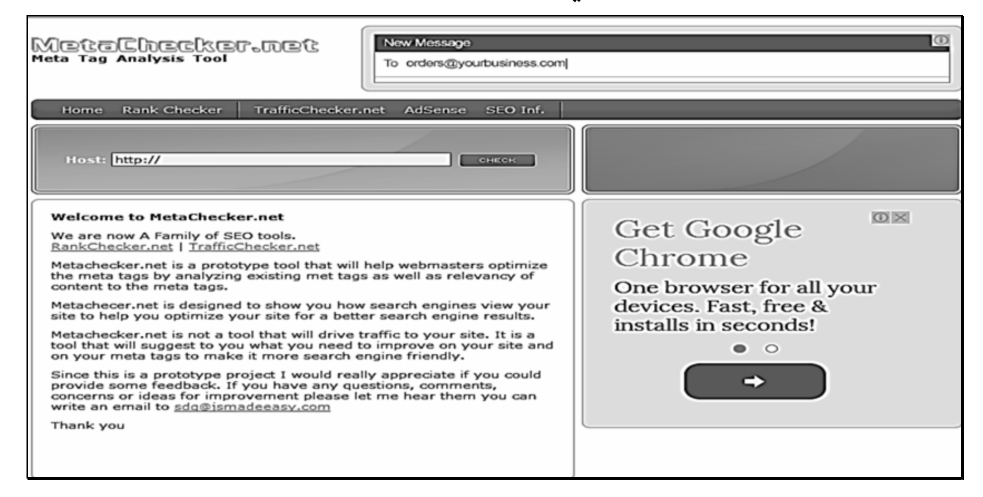
```
<!DOCTYPE html PUBLIC "-//W3C//DTD XHTML 1.0 Strict//EN" "http://www.w3.org/TR/xhtml1/DTD/xhtml1-strict.dtd">
<!-- saved from url=(0160)http://www.webreview.dz/?%D8%A7%D9%84%D8%AA%D8%B4%D8%B1%D9%8A%D8%B9%D8%A7%D8%AA-%D9%88-%D8%A7%D9%84%D8%AA%D9%83%D9%86%D9%88%D9
<html xmlns="http://www.w3.org/1999/xhtml" xml:lang="fr" lang="fr" dir="ltr" class="cye-lm"><head><meta http-equiv="Content-Type" content="text/html; ch
        <title>التشريعات و التكنولوجيا و دورها في دعم و تطوير الأرشيف الوطني - [WEBREVIEW]</title>
                <!-- META DATA -->

        <meta http-equiv="Content-language" content="fr">
        <meta name="language" content="fr">
        <style id="nightModeStyle">html.cye-enabled.cye-nm { -webkit-filter: contrast(68%) brightness(108%) invert(); }html.cye-enabled.cye-nm iframe {
img { -webkit-filter: invert(); }html.cye-enabled.cye-nm i, html.cye-enabled.cye-nm video { -webkit-filter: invert(); }html.cye-enabled.cye-nm i span {
object[type="application/x-shockwave-flash"],html.cye-enabled.cye-nm object[type="application/x-shockwave-flash"] { -webkit-filter: invert(); }html.cye-
flash"],html.cye-enabled.cye-nm embed[type="application/x-shockwave-flash"] { -webkit-filter: invert(); }</style><style id="cyebody">html.cye-enabled.cy
!important;border-color:rgb(156, 105, 205) !important;background-image:none !important;color:#000000  !important}</style><style id="cyediv">html.cye-ena
!important;border-color:rgb(156, 105, 205) !important;background-image:none !important;color:#000000  !important}</style><style id="cyetable">html.cye-e
!important;border-color:rgb(156, 105, 205) !important;background-image:none !important;color:#000000  !important}html.cye-enabled.cye-lm td{background-c
205) !important;background-image:none !important;color:#000000  !important}</style><style id="cyetextInput">html.cye-enabled.cye-lm input[type=text]{bac
color:rgb(156, 105, 205) !important;background-image:none !important;color:#000000  !important}html.cye-enabled.cye-lm textarea{background-color:#738274
!important;background-image:none !important;color:#000000  !important}</style><style id="cyeselect">html.cye-enabled.cye-lm select{background-color:#738
!important;background-image:none !important;color:#000000  !important}</style><style id="cyeul">html.cye-enabled.cye-lm ul{background-color:#738274 !imp
!important;background-image:none !important;color:#000000  !important}</style><style id="cyeChangeByClick">html.cye-enabled.cye-lm .cye-lm-tag,html.cye-
!important;border-color:rgb(156, 105, 205) !important;background-image:none !important;color:#000000  !important}</style><meta http-equiv="Content-Style
        <meta http-equiv="Content-Script-Type" content="text/javascript">
        <meta name="generator" content="SPIP 3.0.17 [21515]">
        <meta name="robots" content="index,follow">
        <link rel="schema.DCTERMS" href="http://purl.org/dc/terms/">
        <link rel="schema.DC" href="http://purl.org/dc/elements/1.1/">

  <!-- META article -->
  <meta name="Description" content="

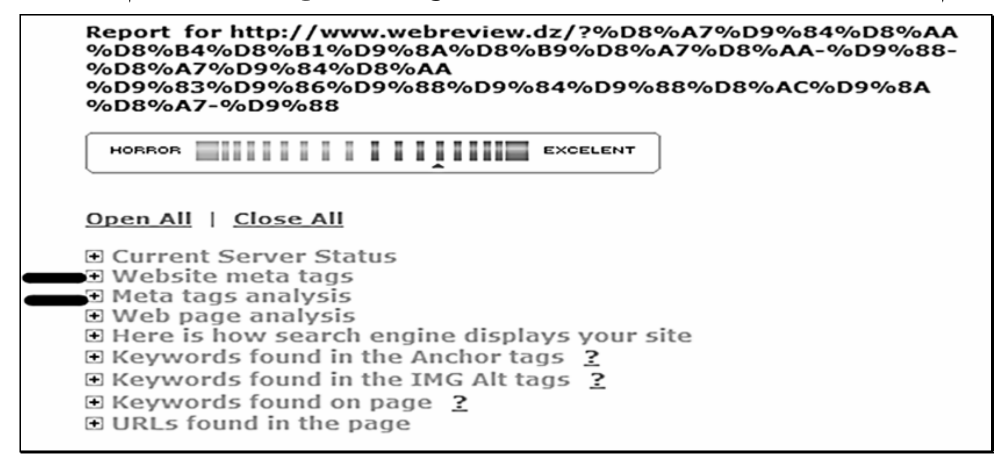
مما يجعل الاهتمام به من عوامل الاهتمام بالتشريعات و التكنولوجيا. فالتشريعات تساعد في التنظيم و الحماية المادية والدعم المادي و المالي, وفي هذا
القانونية و الضرورية لضمان السير الأمثل لقطاع الأرشيف الوطني , و فيما يتعلق بالتكنولوجيا فهي تتيح فرصة التحكم و السيطرة و كذا الاستفادة الجيدة من
المنابع المعرفية و ال ">
  <meta name="Author" content="بودوشة أحمد">
  <meta name="Keywords" content="تكنولوجيا المعلومات , الأرشيف الوطني, التشريعات القانونية, تطوير الأرشيف, الأدوات القانونية">
  <!-- META Dublin Core - voir: http://uk.dublincore.org/documents/dcq-html/  -->
  <meta name="DC.title" content="التشريعات و التكنولوجيا و دورها في دعم و تطوير الأرشيف الوطني">
  <meta name="DC.language" scheme="ISO639-1" content="fr">
  <meta name="DC.identifier" scheme="DCTERMS.URI" content="http://www.webreview.dz/./?%D8%A7%D9%84%D8%AA%D8%B4%D8%B1%D9%8A%D8%B9%D8%A7%D8%AA-%D9%88-
%D8%A7%D9%84%D8%AA%D9%83%D9%86%D9%88%D9%84%D9%88%D8%AC%D9%8A%D8%A7-%D9%88">
  <meta name="DC.source" scheme="DCTERMS.URI" content="http://www.webreview.dz">
  <meta name="DC.creator" content="بودوشة أحمد">
```

الشكل رقم (2-1) شفرة المصدر Source Code لأحد مقالات الدوريات الإلكترونية العربية في مجال المكتبات



الشكل رقم (2-2) برنامج ميتاشيكر MetaChecker.net

ثـم يعـرض البرنـامج في صـفحة مسـتقلة نتيجـة تحليـل عنـاصر مـا وراء البيانـات هـذا بالإضافة إلى عدد آخر من العناصر الخاصة بالموقع مثل حالة الخـادم، والـروابط الموجـودة في الموقع، وتم الاقتصار على التحليل الوارد تحت: "ما وراء البيانات في الموقـع Website Meta Tags" والذي يرد تحته قائمة بعنـاصر مـا وراء البيانـات التـي تتـوافر في هـذا الموقـع، وعـدد تمثيلات قيم كل عـنصر، و"تحليـل تيجـان الميتـا Meta Tags Analysis " والـذي يـرد تحتـه تقييم لعناصر ما وراء البيانات وعلاقة كل منها بالموقع وهذا يتضح من الشكل رقم (2-3).



الشكل رقم (2-3) نتيجة تحليل أحد المقالات على برنامج ميتاشيكر

إلا أن هناك عدد من الملاحظات على هذا البرنامج تشمل التالي:

1- لم يستطع البرنامج تحليل مقالات كل من دورية دراسـات المعلومـات، والمجلـة العربيـة للدراسات المعلوماتية ويقوم بإظهار رسالة خطأ كما في الشكل رقم (2-4).

2- رغم قدرة البرنامج على التعرف على وحصر عناصر مـا وراء البيانـات وفقـا لخطـة بـؤرة دبلن إلا أنه لا يستطيع تقديم تقييم لمحتوى هذه العناصر[1].

(1) جوهري، عزة فاروق. الميتاداتا ودعم استرجاع المحتوى الرقمي للصحف العربية الإلكترونية في البيئة الرقمية: دراسة تطبيقية لمـدى تمثيلهـا في بعـض الصـحف المصريـة والسـعودية. بحـوث في علـم المكتبـات والمعلومات. ع 7 (سبتمبر 2011). ص 75.

```
Server Error in '/' Application.

Object reference not set to an instance of an object.

Description: An unhandled exception occurred during the execution of the current web request. Please review the stack trace for more information about the error and where it originated in the code.

Exception Details: System.NullReferenceException: Object reference not set to an instance of an object.

Source Error:

Line 213:            {
Line 214:                  Metatag cur = new Metatag();
Line 215:                  cur.Content = m.Attributes["content"].Value;
Line 216:                  if (m.Attributes["http-equiv"] != null)
Line 217:                  {

Source File: d:\www\metachecker.net\Site\Analise.aspx.cs    Line: 215

Stack Trace:

[NullReferenceException: Object reference not set to an instance of an object.]
   Analise.AnaliseMetatags(XhtmlDocument document) in d:\www\metachecker.net\Site\Analise.aspx.cs:215
   Analise.Page_Load(Object sender, EventArgs e) in d:\www\metachecker.net\Site\Analise.aspx.cs:112
   System.Web.Util.CalliHelper.EventArgFunctionCaller(IntPtr fp, Object o, Object t, EventArgs e) +24
   System.Web.Util.CalliEventHandlerDelegateProxy.Callback(Object sender, EventArgs e) +41
   System.Web.UI.Control.OnLoad(EventArgs e) +131
   System.Web.UI.Control.LoadRecursive() +65
   System.Web.UI.Page.ProcessRequestMain(Boolean includeStagesBeforeAsyncPoint, Boolean includeStagesAfterAsyncPoint) +2427

Version Information: Microsoft .NET Framework Version:2.0.50727.5485; ASP.NET Version:2.0.50727.5491
```

الشكل رقم (2-4) رسالة خطأ من برنامج ميتاشيكر عند إدخال الموقع URL لأحد مقالات دورية المجلة العربية للدراسات المعلوماتية

1-2 مدى توافر عناصر ما وراء البيانات في مقـالات الـدوريات الإلكترونيـة العربيـة في مجـال المكتبات والمعلومات:

يتضح من الجدول رقم (2-3) أن جميع مقالات الدوريات الإلكترونية العربية في مجـال المكتبات والمعلومات (222 مقالة) وردت بها عناصر ما وراء البيانات في جزء الرأس head من شفرة مصدر source code هذه المقالات، وذلك بخلاف عنصر (تاج) العنوان الذي يجـب أن يتوافر في أي صفحة تتاح على العنكبوتية، وعادة ما يتم إنشاءه آليـا بواسـطة بـرامج تصـميم المواقع أو برامج إدارة محتوى المواقع CMS.

كما تتأكد هذه النتيجة على مستوى الدورية الواحدة ؛ حيث كانت نسبة المقالات التي وردت بها ما وراء البيانات في كل دورية على حدة 100%.

الجدول رقم (2-3) مدى توافر ما وراء البيانات في مقالات الدوريات الإلكترونية العربية في المجال

الدورية	توجد عناصر ما وراء البيانات		لا توجد عناصر ما وراء البيانات	
	العدد	النسبة من مقالات كل دورية	العدد	النسبة من مقالات كل دورية
Cybrarians Journal	108	%100	صفر	صفر%
العربية 3000	51	%100	صفر	صفر%
دراسات المعلومات	42	%100	صفر	صفر%
المجلة العربية للدراسات المعلوماتية	15	%100	صفر	صفر%
مجلة المكتبات والمعلومات	6	%100	صفر	صفر%
الإجمالي	222	%100	صفر	صفر%

2-2 خطط ما وراء البيانات المستخدمة في مقالات الدوريات الإلكترونية العربية محل الدراسة:

من أبرز خطط ما وراء البيانات في بيئة العنكبوتية - كما ورد تفصيلا في الفصل الأول - هي: عناصر ما وراء البيانات كما نصت عليها لغة ترميز النص الفائق HTML Metadata، ومعيار بؤرة دبلن، بالإضافة إلى الخطط التي ظهرت في الفترة الأخيرة وتدخل في إطار بيئة العنكبوتية الدلالية وتشمل كل من بروتوكول الرسم البياني (og)، وبطاقات تويتر Twitter Cards. ويعرض الجدول رقم (2-4) خطط ما وراء البيانات التي استخدمت مع عناصر ما وراء البيانات الواردة في مقالات الدوريات الإلكترونية محل الدراسة.

يتضح من الجدول رقم (2-4) ما يلي:

- جميع الدوريات الإلكترونية العربية في مجال المكتبات والمعلومات استخدمت تيجان ما وراء البيانات التي نصت عليها لغة ترميز النص الفائق HTML، والتي عادة ما تعرف بتيجان الميتا HTML Meta Tags. وهذا يتوافق مع ما أشارت إليه دراسة سابقة من

الجدول رقم (2-4) خطط ما وراء البيانات المستخدمة في مقالات الدوريات الإلكترونية العربية في المجال

خطط ما وراء البيانات / الدورية	عناصر ما وراء بيانات لغة ترميز النص الفائق HTML Metadata		بؤرة دبلن	
	العدد	النسبة من مقالات كل دورية	العدد	النسبة من مقالات كل دورية
Cybrarians Journal	108	%100	صفر	صفر%
العربية 3000	51	%100	صفر	صفر%
دراسات المعلومات	42	%100	صفر	صفر%
المجلة العربية للدراسات المعلوماتية	15	%100	صفر	صفر%
مجلة المكتبات والمعلومات	6	%100	6	%100
الإجمالي	222	%100	6	%2.70

أن أكثر الخطط استخداما على الإنترنت هي تيجان الميتا للغة ترميز النص الفائق HTML Meta Tags.[1]

- استخدمت دورية واحدة فقط هي دورية مجلة المكتبات والمعلومات معيار بؤرة دبلن لترميز عناصر ما وراء البيانات الواردة بمقالاتها هذا بجانب استخدامها لتيجان الميتا للغة ترميز النص الفائق HTML Meta Tags المستخدمة في الدوريات الأخرى، أي أن بؤرة دبلن تم استخدامها في 6 مقالات فقط – جميعها مقالات مجلة المكتبات والمعلومات - بنسبة تقدر بنحو2.70 % من إجمالي المقالات عينة الدراسة. وتؤكد هذه النتيجة على عدم اهتمام الدوريات الإلكترونية العربية في مجال المكتبات والمعلومات باستخدام بؤرة دبلن لترميز عناصر ما وراء البيانات، وقد يرجع هذا إلى عدة أسباب:

(1) Zhang, Jin, jastram, Iris. A Study of Metadata Element Co-occurrence. Online Information Review. Vol. 30, No. 4 (2006). P 428.

- بصفة عامة، في بيئة العنكبوتية لا يوجد إلزام بإدراج عناصر ما وراء البيانات في الصفحات والمواقع المتاحة عليها، فهذا يخضع بالدرجة الأولى لرغبة وقدرة منشئي هذه المواقع على إدراج هذه العناصر. وبالتالي ليس هناك إلزام باستخدام خطة بعينها، كما أن إنشاء المواقع عادة ما يعهد به إلى متخصصين في الحاسبات وتصميم المواقع، وهم في الأغلب ليسوا على دراية بمعيار بؤرة دبلن.
- لم توصي أي من محركات البحث الرئيسية مثل جوجل، أو ياهو، أو بينج بضرورة إدراج عناصر ما وراء البيانات وفقا لمعيار بؤرة دبلن[1].

• تتفق نتيجة عدم اهتمام دوريات المكتبات والمعلومات الإلكترونية العربية في مجال المكتبات والمعلومات باستخدام معيار بؤرة دبلن مع عدد من الدراسات السابقة ؛ حيث أشارت إلى نفس النتيجة دراسة محمد أمين مرغلاني، وسوزان مصطفى فلمبان عن ما وراء البيانات في المواقع الإلكترونية للمكتبات الجامعية السعودية حيث أفادت بعدم استخدام معيار بؤرة دبلن في إنشاء ما وراء البيانات لهذه المواقع[2]. وكذلك دراسة سوزان مصطفى فلمبان والتي توصلت إلى عدم الاهتمام بتطبيق معيار بؤرة دبلن في إعداد عناصر ما وراء بيانات مدونات المكتبات الجامعية، حيث لم يرد إلا في مدونة واحدة فقط هي مدونة What's New @ UWM Libraries التابعة لمكتبات جامعة ميلواكي – ويسكنسن بالولايات المتحدة الأمريكية وفي عنصر العنوان فقط[3]، كما أشارت إلى نفس النتيجة دراسة بدوية محمد البسيوني عن مواقع التمييز

(1) Meta tags that Google understands - Search Console Help. [online] Available at: https://support.google.com/webmasters/answer/79812?hl=en .[Accessed 27/5/2014].

(2) مرغلاني، محمد أمين، فلمبان، سوزان مصطفى. الميتاداتا في المواقع الإلكترونية للمكتبات الجامعية السعودية [على الخط المباشر]. دراسات المعلومات. ع 2 (مايو 2008). ص ص 7 – 46. متاح على: http://journals.psiscs.com.sa/index.php/ijs/article/view/179/147 . تاريخ الإتاحة [2015/3/14].

(3) فلمبان، سوزان بنت مصطفى بن عباس. مدونات المكتبات الجامعية: دراسة تحليلية في دعم برمجيات المدونات للميتاداتا (واصفات البيانات) / إعداد سوزان بنت مصطفى بن عباس فلمبان؛ إشراف محمد أمين عبد الصمد مرغلاني، بدوية محمد البسيوني. أطروحة (دكتوراه). جامعة الملك عبد العزيز، كلية الآداب والعلوم الإنسانية، قسم المعلومات، 2011.

الرقمي العربية على الإنترنت حيث أبرزت قلة عدد المواقع المعتمدة على معيار بؤرة دبلن حيث لم تظهر تيجان بؤرة دبلن إلا في ثلاثة مواقع بنسبة تمثل 1.15% من إجمالي المواقع موضوع الدراسة وهي نسبة ضعيفة[1].

3-2 التوزيع العددي لعناصر (تيجان) ما وراء البيانات في مقالات الدوريات الإلكترونية العربية محل الدراسة:

يبين الجدول رقم (5-2) التوزيع العددي لعناصر ما وراء البيانات الواردة في مقالات الدوريات الإلكترونية محل الدراسة، مع ملاحظة أنه في حالة تكرار العنصر أو حتى في حالة وروده باستخدام خطط ما وراء بيانات مختلفة تم حسابه كعنصر واحد لمرة واحدة، على سبيل المثال في مجلة المكتبات والمعلومات – كما اتضح سابقا –يتم استخدام تيجان الميتا بجانب معيار بؤرة دبلن في إدراج عناصر ما وراء البيانات وعلى هذا الأساس تكررت بعض العناصر مثل عنصر (تاج) المؤلف حيث ورد مرة بصيغة تيجان الميتا وورد مرة أخرى وفقا لبؤرة دبلن كالتالي:

```
<meta name="Author" content="بودوشة أحمد">
...
<meta name="DC.creator" content="بودوشة أحمد">
```

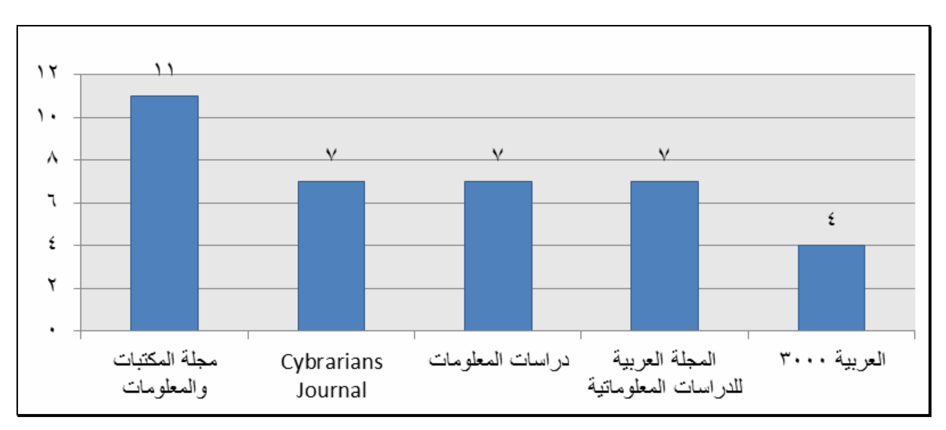
(1) البسيوني، بدوية محمد. تيجان الميتاداتا meta tags ومدى تمثيلها في صفحات الويب: دراسة تطبيقية على مواقع التميز الرقمي العربية على الانترنت. الاتجاهات الحديثة في المكتبات والمعلومات. مج 16، ع32 (يوليو 2009) ص ص 179 – 224.

الجدول رقم (2-5) التوزيع العددي لعناصر ما وراء البيانات في مقالات الدوريات الإلكترونية العربية في مجال المكتبات والمعلومات

الدورية	عدد المقالات	عدد عناصر ما وراء البيانات الواردة بكل مقالة	إجمالي عدد عناصر ما وراء البيانات في إجمالي مقالات كل دورية
Cybrarians Journal	108	7	756
العربية 3000	51	4	204
دراسات المعلومات	42	7	294
المجلة العربية للدراسات المعلوماتية	15	7	105
مجلة المكتبات والمعلومات	6	11	66
الإجمالي	222	36	1425

يتضح من الجدول رقم (2-5) ما يلي:

- بلغ إجمالي عدد عناصر (تيجان) ما وراء البيانات في مقالات الدوريات الإلكترونية في مجال المكتبات والمعلومات محل الدراسة (1425) عنصرا.
- تجدر الإشارة إلى ملاحظة هامة وهي ثبات عدد عناصر ما وراء البيانات في مقالات كل دورية على حدة، ولم تُظهر أي دورية اختلافا في عدد عناصر ما وراء البيانات بين مقالاتها. وفي هذا إشارة إلى أن عناصر ما وراء البيانات يتم إدراجها آليا ولا يتم ترميزها أو تكويدها يدويا، إنما يتم الاعتماد في ذلك على برامج التصميم وإدارة محتوى المواقع التي تستخدمها هذه الدوريات، ولا يتم التدخل البشري إلا بإدخال القيم الخاصة بهذه العناصر. كما أن هذه الملاحظة تشير إلى أن كل دورية تنتهج منهجا ثابتا في إدراج عناصر ما وراء البيانات الخاصة بها سواء من حيث العدد، أو الأنواع المستخدمة، أو طريقة وأسلوب الترميز وذلك في جميع مقالاتها.وعلى أثر هذه الملاحظة فإنه في كثير من الأحيان عند التحليل الإحصائي، تم تفضيل عرض نتائج تحليل المقالات وفقا لكل دورية وليس بشكل منفصل كمقالات مستقلة، وهذا ما سيتأكد في نقاط التحليل التالية.



الشكل رقم (2-5) عدد تيجان ما وراء البيانات في مقالات كل دورية

- بلغ أقصى عدد من عناصر ما وراء البيانات التي وردت في مقالات الدوريات الإلكترونية محل الدراسة (11) عنصرا،وذلك في مقالات مجلة المكتبات والمعلومات، بينما ورد عدد سبعة (7) عناصر في مقالات كل من دورية Cybrarians Journal، ودراسات المعلومات، والمجلة العربية للدراسات المعلوماتية.
- ورد أقل عدد من عناصر ما وراء البيانات في مقالات العربية 3000، حيث ورد بها أربعة (4) عناصر فقط لما وراء البيانات.
- بلغ متوسط عدد عناصر ما وراء البيانات التي تستخدمها كل دورية في مقالاتها نحو 7.2 عنصرا[*]، وهذا المتوسط يقترب مما يتواجد في المواقع الصحفية العربية حيث بلغ نحو 7.6 عنصرا[1]، بينما كان هذا المتوسط أكبر مما توافر في مواقع دوريات الوصول الحر في مجال المكتبات والمعلومات والذي بلغ 3.25 عنصرا لكل دورية [2].

(*) تم حساب متوسط عدد العناصر التي تستخدمها كل دورية في مقالاتها بقسمة إجمالي العناصر 36 على عدد الدوريات الخمسة.

(1) جوهري، عزة فاروق.مصدر سابق. ص 75.

(2) العربي، أحمد عبادة. الميتاداتا ودورها في دعم المحتوى الرقمي: دراسة تطبيقية على مواقع دوريات الوصول الحر في المكتبات والمعلومات على الإنترنت. مجلة مكتبة الملك فهد الوطنية. مج 16، ع 1 (المحرم / جمادي الآخرة 1431 هـ / يونيو 2010). ص 140.

4-2 التوزيع النوعي لعناصر ما وراء البيانات في مقالات الدوريات الإلكترونية محل الدراسة:

كما ورد سابقا عن ثبات أعداد عناصر ما وراء البيانات في مقالات كل دورية على حدة، يتضح أيضا ثبات أنواع عناصر ما وراء البيانات المستخدمة في مقالات كل دورية؛ حيث استقرت كل دورية على استخدام نفس أنواع عناصر ما وراء البيانات في جميع المقالات التي تتيحها على موقعها، وبالتالي إذا كانت دورية ما تستخدم عنصر الوصف على سبيل المثال، فإن هذا العنصر يظهر في جميع مقالاتها بنسبة 100%، وفي هذا تأكيد على ما سبق الإشارة إليه من أن هذه العناصر يتم إدراجها آليا. ووفقا لذلك فقد تم توزيع أنواع هذه العناصر وفقا لكل دورية.

ووزعت أنواع عناصر ما وراء البيانات تحت فئتين رئيسيتين هما: أ- فئة عناصر ما وراء البيانات الوصفية وهي العناصر التي تهتم بوصف محتوى المقالة مثل العنوان title، والوصف description، والكلمات المفتاحية keywords. وعادة ما ترد هذه العناصر في الخاصية name لعنصر أو تاج الميتا meta tag. أما الفئة الثانية فهي عناصر ما وراء البيانات الفنية أو التقنية التي تهتم بالجوانب الفنية الخاصة بالصفحة مثل مجموعة التمثيلات المستخدمة charset، أو نوع محتوى الصفحة text/html وعادة ما ترد هذه العناصر في الخاصية http-equiv لتاج الميتا. ويعرض الجدول التالي رقم (2-6) أنواع عناصر ما وراء البيانات المستخدمة في مقالات الدوريات الإلكترونية العربية في المجال مرتبة تنازليا وفقا لمعدل ورودها في مقالات هذه الدوريات.

يتضح من الجدول رقم (2-6) ما يلي:

- بلغ إجمالي أنواع عناصر ما وراء البيانات المستخدمة في إجمالي مقالات الدوريات الإلكترونية العربية في المجال (13) عنصرا هي: العنوان title، والوصف description، والكلمات المفتاحية keywords /أو الموضوع subject، والمؤلف author/ أو المنشيء creator، والتاريخ date، واللغة language، والمحدد identifier، والمصدر source، ونوع المحتوى content-type، والمولد generator، والروبوت robot، والعرض viewport، ومدى التوافق X-UA-compatible.

الجدول رقم (2-6) التوزيع النوعي لعناصر ما وراء البيانات في مقالات الدوريات الإلكترونية العربية في مجال المكتبات والمعلومات

أنواع عناصر ما وراء البيانات	مقالات Cybrarians Journal (108 مقالة)	مقالات العربية 3000 (51 مقالة)	مقالات دراسات المعلومات (42 مقالة)	مقالات المجلة العربية للدراسات المعلوماتية (15 مقالة)	مقالات مجلة المكتبات والمعلومات (6 مقالة)	عدد الدوريات التي استخدمت هذا العنصر	النسبة من إجمالي عدد الدوريات (5 دوريات)	إجمالي عدد مرات ظهور العنصر	النسبة من إجمالي المقالات (222)
أ- فئة عناصر ما وراء البيانات الوصفية									
العنوان title	✓	✓	✓	✓	✓	5	%100	222	%100
الوصف description	✓	✓	✓	✓	✓	5	%100	222	%100
الكلمات المفتاحية / الموضوع keywords / subject	✓	✓	✓	✓	✓	5	%100	222	%100
المؤلف / المنشيء author /creator	✓	✗	✗	✗	✓	2	%40	114	%51.35
التاريخ date	✗	✗	✗	✗	✓	1	%20	6	%2.70
اللغة language	✗	✗	✗	✗	✓	1	%20	6	%2.70
المحدد identifier	✗	✗	✗	✗	✓	1	%20	6	%2.70
المصدر source	✗	✗	✗	✗	✓	1	%20	6	%2.70

أنواع عناصر ما وراء البيانات	مقالات Cybrarians Journal (108 مقالة)	مقالات العربية 3000 (51 مقالة)	مقالات دراسات المعلومات (42 مقالة)	مقالات المجلة العربية للدراسات المعلوماتية (15 مقالة)	مقالات مجلة المكتبات والمعلومات (6 مقالة)	عدد الدوريات التي استخدمت هذا العنصر	النسبة من إجمالي عدد الدوريات (5 دوريات)	إجمالي عدد مرات ظهور العنصر	النسبة من إجمالي المقالات (222)
ب- فئة عناصر ما وراء البيانات الفنية									
نـــــوع المحتـــــوى content-type	✓	✓	✓	✓	✓	5	%100	222	%100
المولد generator	✓	✗	✓	✓	✓	4	%80	170	%76.58
الروبوت robots	✓	✗	✗	✗	✓	2	%40	114	%51.35
العرض viewport	✗	✗	✓	✓	✗	2	%40	57	%33.78
X-UA-Compatible	✗	✗	✓	✓	✗	2	%40	57	%33.78

- توزعت أنواع عناصر ما وراء البيانات على فئتين أساسيتين هما: عناصر ما وراء البيانات الوصفية، وعناصر ما وراء البيانات الفنية، وانتمى العدد الأكبر من عناصر ما وراء البيانات المدرجة في مقالات الدوريات الإلكترونية العربية في المجال إلى فئة عناصر ما وراء البيانات الوصفية حيث بلغ عددها ثمانية (8) عناصر شملت العنوان، والوصف، والكلمات المفتاحية، والمؤلف، والتاريخ، واللغة، والمحدد، والمصدر. بينما ورد في فئة عناصر ما وراء البيانات الفنية خمسة (5) عناصر هي: نوع المحتوى، والمولد، والروبوت، والعرض، ومدى التوافق.
- كانت أكثر العناصر (التيجان) ورودا في مقالات الدوريات الإلكترونية العربية في المجال كل من: العنوان، والوصف، والكلمات المفتاحية، ونوع المحتوى حيث ورد كل منها في جميع مقالات الدوريات الإلكترونية العربية محل الدراسة (بنسبة 100%)، ويعد تاج العنوان والوصف من أبرز عناصر ما وراء البيانات التي عادة ما توصي محركات البحث بإدارجها في الصفحات والمواقع.
- جاء في المرتبة الثانية عنصر المولد generator ويقصد به البرنامج المستخدم في إنشاء وتوليد الصفحة أو الموقع[1]، وقد ورد هذا العنصر في 170 مقالة بنسبة حوالي 76.58% من إجمالي المقالات، وقد استخدمت جميع الدوريات محل الدراسة هذا العنصر فيما عدا دورية العربية 3000 (أي أربع دوريات)، ويرجع السبب في ارتفاع معدل ظهور هذا العنصر هو أنه عادة ما يتم إدراجه آليا في حالة تصميم المواقع باستخدام برامج التصميم أو برامج إدارة محتوى المواقع، ويؤكد ذلك قيم هذا التاج حيث كانت قيمة هذا التاج في دورية Cybrarians Journal برنامج جوملا الإصدارة Joomla! 1.5 1.15 وهو أحد برامج إدارة محتوى المواقع الشائعة الاستخدام كما يلي:

(1) W3C. Document metadata — HTML5 [online].Available at: http://www.w3.org/TR/html5/document-metadata.html#the-meta-elemen . Accessed [12/10/2015].

<meta name="generator" content="Joomla! 1.5 - Open Source Content Management">

بينما كانت قيمة هذا التاج في كل من دورية دراسات المعلومات والمجلة العربية للدراسات المعلوماتية برنامج المجلات العلمية 2.4.3.0 ويبدو أنه أحد البرامج المحلية التي تستخدمها هذه الدوريات.

أما مجلة المكتبات والمعلومات فكانت قيمة هذا التاج أحد برامج إدارة محتوى المواقع وهو برنامج Système de Publication pour l'Internet (SPIP 3.0.17).

- جاء في المرتبة الثالثة كل من عنصر المؤلف وعنصر الروبوت الذي استخدمته دوريتان في مقالاتها وهي كل من Cybrarians Journal، ومجلة المكتبات والمعلومات، حيث ورد كل منهما في إجمالي 114 مقالة بنسبة بلغت 51.35%. ومن الجدير بالذكر أن عنصر الروبوت يحدد فيه كيفية تعامل زواحف أو عناكب محركات البحث مع الصفحة، وفي حالة عدم وجوده فإن الوضع الافتراضي هو أن تقوم محركات البحث بتكشيف الصفحة ومتابعتها[1]، و قد ورد هذا العنصر في الدوريات التي ورد بها كالتالي:

<meta name="robots" content="index, follow">

وبالنظر إلى قيم هذا التاج والتي كانت التكشيف والتتبع index,follow وهي قيم تتفق مع الوضع الافتراضي لتعامل زواحف محركات البحث مع الصفحة، فإن إدراج هذا العنصر في هذه الدوريات يعد أمرا زائدا عن الحاجة. هذا بالإضافة إلى أنه عادة ما تقوم المواقع والصفحات بإدراج ملف يسمى Robots.txt في ملفات الجذر root للموقع، وذلك لتحديد كيفية تعامل محركات البحث مع الموقع وصفحاته.

- جاء في المرتبة الرابعة اثنان من العناصر التي تنتمي إلى ما وراء البيانات الفنية وهما عنصر العرض viewport الخاص بطريقة عرض المقالات على أجهزة المحمول، وعنصر X-UA-Compatible لمدى التوافق مع متصفح إنترنت أكسبلورر Internet

(1) Meta tags that Google understands - Search Console Help. [online]. Op. Cit.

Explorer، حيث ورد هذان العنصران في جميع مقالات اثنان من الدوريات محل الدراسة وهما دراسات المعلومات، والمجلة العربية للدراسات المعلوماتية، وذلك في عدد 57 مقالة بنسبة حوالي 33.78% من إجمالي المقالات.

- كانت أقل العناصر ورودا في مقالات الدوريات الإلكترونية العربية في المجال العناصر التالية: اللغة، والتاريخ، والمحدد، والمصدر، حيث وردت هذه العناصر في مقالات دورية واحدة فقط (أي 6 مقالات) هي مجلة المكتبات والمعلومات بنسبة 2.70%.

5-2 جودة ما وراء البيانات في مقالات الدوريات الإلكترونية العربية في مجال المكتبات والمعلومات:

تعرف الجودة بصفة عامة بأنها مدى الملاءمة للغرض أو الهدف fitness for purpose [1]، ووفقا لذلك يمكن القول بأن جودة ما وراء البيانات هي مدى ملاءمة ما وراء البيانات للغرض أو الهدف المقصود منها، كما يمكن تعريفها بأنها الحكم الموضوعي على مدى صلاحية وفائدة ما وراء البيانات[2]. فليست العبرة بتوافر عناصر ما وراء البيانات فحسب، ولكن تكمن أهمية هذه العناصر في قدرتها على وصف المصادر الإلكترونية من خلال ما يسجل بها من قيم. ومما لا شك فيه أن توافر ما وراء بيانات بمستوى جيد يؤثر على تأدية وظائفها المرتبطة بالعثور على المصادر واكتشافها.

---

(1) Palavitsinis, Nikos. Metadata Quality Issues in Learning Repositories [online].2013.P.57. Available at: http://blog.agroknow.com/wp-content/uploads/2014/02/Palavitsinis_Thesis_Final221113.pdf . Accessed [5/5/2016].

(2) Hilligoss, B., & Rieh, S.Y. Developing a unifying framework of credibility assessment: Construct, heuristics, and interaction in context. Information Processing & management. Vol. 44, No. 4 (2008), 1467–1484. As cited in: Palavitsinis, Nikos. Metadata Quality Issues in Learning Repositories [online] / supervisors Nikos Manouselis, Salvador Sánchez-Alonso.(Phd). Universidad de Alcalá, Departamento de Ciencias de la Computación.2013. p. 58. Available at: http://blog.agroknow.com/wp-content/uploads/2014/02/Palavitsinis_Thesis_Final221113.pdf. Accessed [5/5/2016].

وهناك العديد من المعايير التي تقاس بها جودة ما وراء البيانات، إلا أنها عادة ما تطبق على ما وراء البيانات في فهارس المكتبات والمستودعات الرقمية، ومنها على سبيل المثال لا الحصر ما يلي: - الاكتمال Completeness بمعنى توافر عناصر البيانات القادرة على وصف المصدر على أكمل وجه ممكن وأن جميع هذه العناصر مكتملة، - الصحة أو الدقة Accuracy ويقصد بها مدى صحة ودقة البيانات وخلوها من الأخطاء في عناصر ما وراء البيانات سواء الأخطاء الكتابية أوالأخطاء الحقائقية، - المطابقة للتوقعات Conformance to Expectations أي الدرجة التي تلبي فيها ما وراء البيانات احتياجات مجتمع ما من المستفيدين في القيام بوظيفة ما، - الاتساق والترابط المنطقي بين عناصر ما وراء البيانات Logical Consistency and Coherence، - الإتاحة Accessibility فما وراء البيانات يجب أن تكون متاحة وقابلة للقراءة والفهم، - مناسبة لكل وقت Timeliness فما وراء البيانات يجب أن تتغير إذا ما تغير المصدر الموصوف، - مصدر ما وراء البيانات provenance أي الجهة المسؤولة عن إعداد ما وراء البيانات [(1)].

وهذه المعايير من السهل تطبيقها لقياس جودة ما وراء البيانات في بيئة المكتبات والمستودعات الرقمية لأنها بيئة مغلقة، يتم فيها الالتزام بخطة عمل يمكن على أساسها قياس هذه المعايير، أما في بيئة العنكبوتية فيصعب تطبيق مثل هذه المعايير، ولذلك فإن هذه الدراسة جعلت من مدى صلة القيم الواردة في ما وراء البيانات بمحتوى كل مقالة على حدة، ودقتها في وصفها معيارا لقياس جودة ما وراء البيانات، وهو ما يمكن اعتباره جزءا من معيار الصحة أو الدقة Accuracy.

ومن بين أنواع العناصر الثلاثة عشر (13) التي وردت في مقالات الدوريات

(1) Ochoa, Xavier. Metadata Quality [online]. Pp.67-68. In: Handbook of Metadata, Semantics and Ontologies. edited by Miguel-Angel Sicilia. New Jersey: World Scientific, 2014.Available at: https://books.google.com.eg/books?id=mgS3CgAAQBAJ&pg=PA66&lpg=PA66&dq=ahUKEwiqvKul-sHMAhVhIcAKHY3zBRkQ6AEIPzAJ#v=onepage&q&f=false . Accessed [4/5/2016].

الإلكترونية العربية في المجال، فقد تم قياس الجودة لعناصر ما وراء البيانات التي تنتمي إلى فئة عناصر ما وراء البيانات الوصفية دون غيرها والتي تشمل ثمانية عناصر هي: العنوان title، والكلمات المفتاحية / الموضوع subject / keywords، والوصف description، والمؤلف/ المنشيء author /creator، والتاريخ date، واللغة language، والمحدد identifier، والمصدر source؛ فهذه العناصر هي التي تميز كل مقالة عن الأخرى وتعمل على وصفها. ولم يتم النظر إلى جودة عناصر ما وراء البيانات الفنية لأنها غالبا ما يتم إنشاءها آليا، كما أنها تؤثر على طريقة عرض الموقع والتعامل معه، وليس على كيفية العثور على الموقع واكتشافه.

وقد تم قياس مدى صلة القيم الواردة في أي عنصر من العناصر الثمانية على مقياس ما بين واحد والذي يعبر عن أن القيمة الواردة في هذا العنصر لها صلة بمحتوى المقالة وتعبر عنها بدقة، وصفر أي أن القيمة لا صلة لها بمحتوى المقالة.

1-5-2 عنصر العنوان:

يعد عنصر العنوان من أهم المعلومات التي ينبغي توافرها لصفحة العنكبوتية، حيث يصفه اتحاد العنكبوتية W3C بأنه عنصر مطلوب وضروري، وتدعمه جميع محركات البحث حيث تقوم معظم المحركات باستخدام النص المسجل في تاج العنوان في صفحة النتائج[1] ويوفر هذا العنصر لمحة سريعة عن محتوى صفحة ما ومدى صلتها بأحد الاستفسارات بما يؤثر على اختيار المستفيد لهذه الصفحة[2]. لذلك يُوصى بضرورة التأكد من توافر عنوان يناسب كل صفحة من صفحات الموقع، وأن يكون هذا العنوان معبرا ومختصرا[3].

(1) البسيوني، بدوية محمد. مصدر سابق، ص 191 – 192.

(2) Review your page titles and snippets - Search Console Help [online]. Available at: https://support.google.com/webmasters/answer/35624?rd=1 . [Accessed 27/5/2014]

(3) Ibid.

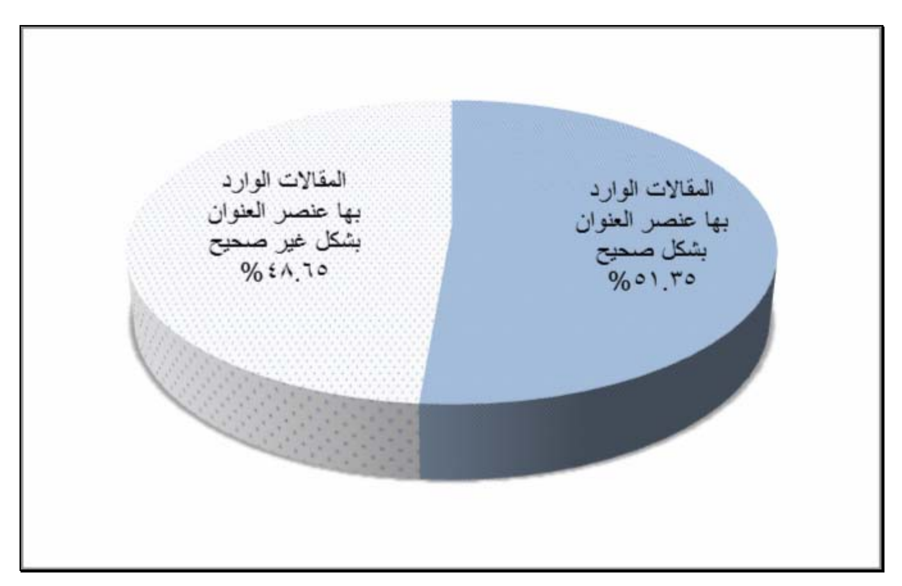
ويوضح الجدول التالي رقم (2-7) جودة عنصر العنوان في مقالات الدوريات الإلكترونية العربية محل الدراسة.

الجدول رقم (2-7) جودة عنصر العنوان في مقالات الدوريات الإلكترونية العربية في مجال المكتبات والمعلومات

الدورية	عدد المقالات	تعبر القيم في عنصر العنوان عن المقالة	النسبة	لا تعبر القيم في عنصر العنوان عن المقالة	النسبة
Cybrarians Journal	108	108	%100	صفر	صفر%
العربية 3000	51	صفر	صفر%	51	%100
دراسات المعلومات	42	صفر	صفر%	42	%100
المجلة العربية للدراسات المعلوماتية	15	صفر	صفر%	15	%100
مجلة المكتبات والمعلومات	6	6	%100	صفر	صفر%
الإجمالي	222	114	%51.35	108	%48.65

يتضح من الجدول السابق رقم (2-7) ما يلي

- بلغ عدد المقالات التي ورد في عنصر (تاج) العنوان الخاص بها قيم صحيحة تعبر عن عناوين هذه المقالات (114) مقالة بنسبة بلغت حوالي 51.35%، بينما بلغ عدد المقالات التي ورد في تاج العنوان الخاص بها قيم لا صلة لها بعناوين هذه المقالات (108) مقالة بما يقدر بنحو 48.65%. وعلى الرغم من أن في هذا إشارة إلى أن نسبة المقالات التي ورد بها عنصر العنوان بشكل صحيح كانت أكبر عن تلك التي ورد بها عنصر العنوان بشكل غير صحيح، إلا أن نسبة تقترب من نصف المقالات لا يرد فيها عنصر العنوان بشكل صحيح مع الأخذ في الاعتبار أنه من أكثر العناصر أهمية.



الشكل رقم (2-6) جودة عنصر العنوان في مقالات الدوريات الإلكترونية العربية في مجال المكتبات والمعلومات

- كانت المقالات التي ورد بها عنصر العنوان بصورة صحيحة تنتمي إلى كل من دورية Cybrarians Journal، ومجلة المكتبات والمعلومات العربية. حيث حرصت كل منهما في جميع مقالاتها على إدراج عنصر العنوان الذي يصف المقالة، فقد حرصت دورية Cybrarians Journal في جميع مقالاتها على تسجيل عنصر العنوان بما يتناسب مع كل مقالة على حدة، كما لوحظ أن هذه الدورية تتبع قواعد الفهرسة الأنجلو أمريكية في تسجيل عنصر العنوان حيث يرد العنوان مصحوبا ببيانات المسئولية بنفس علامات الترقيم التي تنص عليها قواعد الفهرسة كما يتبين من المثال التالي: " اتجاهات موظفي المؤسسة العامة للتأمينات الاجتماعية بمدينة جدة: نحو تطبيق نظام التعاملات الإلكترونية. أ. د. محمد أمين بن عبد الصمد مرغلاني "، وعلى الرغم من أن المفترض تسجيل العنوان فقط في هذا التاج إلا أن تسجيل بيانات المسؤولية أيضا قد يكون نوعا من محاولة إبراز اسم الشخص المسؤول عن العمل، وجعله قابلا للبحث والاسترجاع بصورة أكبر؛ ذلك أن العديد من محركات البحث تستخدم عنصر العنوان بشكل

أساسي في عمليات الاسترجاع. كذلك حرصت مجلة المكتبات والمعلومات على تسجيل عنصر العنوان بشكل صحيح لجميع مقالاتها، سواء الوارد باستخدام تيجان الميتا أو باستخدام بؤرة دبلن، إلا أنها كانت في تيجان الميتا مع عنوان كل مقالة تقوم بإضافة كلمة WEBREVIEW وهو عنوان البوابة التي تتوافر عليها هذه الدورية.

- ظهر عنصر العنوان بصورة غير صحيحة في مقالات كل من الدوريات التالية: العربية 3000، ودراسات المعلومات، والمجلة العربية للدراسات المعلوماتية. ففي العربية 3000 كان العنوان المسجل في جميع مقالات هذه الدورية هو "نادي الإحياء العربي"، وهو اسم النادي المسؤول عن إصدار الدورية، والذي يتاح من خلال موقعه هذه الدورية. أما كل من دراسات المعلومات والمجلة العربية للدراسات المعلوماتية ففي جميع المقالات الصادرة عنهما كان العنوان لجميع المقالات هو المجلات العلمية، وهو ما لا يعبر حتى عن عنوان الدورية على أقل تقدير.

2-5-2 عنصر الوصف:

يسمح عنصر الوصف بتسجيل وصف مختصر لمحتوى الصفحة، وعادة ما يتم إدراج عنصر الوصف بعد عنصر العنوان وقبل عنصر الكلمات المفتاحية وإن كانت لا توجد أهمية فعلية لترتيب العناصر[1]. وقد يستخدم محرك البحث جوجل الوصف الوارد في هذا العنصر في عرض الموقع في صفحة نتائج البحث فيما يعرف بالمتقطف الغني Rich Snippet كما في الشكل التالي:

Google Video
Search and browse all kinds of videos, hosted on sites all over the web, including **Google**, YouTube, MySpace, MetaCafe, GoFish, Vimeo, Biku, and Yahoo **Video**.
video.google.com/ - 108k - Cached - Similar pages - Note this

الشكل رقم (2-7) المقتطف الغني Rich Snippet لمحرك البحث جوجل

(1) البسيوني، بدوية محمد. مصدر سابق، ص 193.

ويُوصى في إعداد هذا العنصر بإعداد وصف مختلف يصف بشكل صحيح كل صفحة من صفحات الموقع، ولا يعد استخدام نفس الوصف لكل صفحة من صفحات الموقع أمرا مساعدا عندما تعرض الصفحات المستقلة في نتائج البحث. وفي هذه الحالات فإن محرك البحث جوجل يميل أكثر إلى عدم عرض هذا النص المتكرر[1]، ويقوم آليا بإعداد وصف آخر.

ويوضح الجدول رقم (2-8) جودة قيم عنصر الوصف في مقالات الدوريات الإلكترونية العربية في مجال المكتبات والمعلومات.

الجدول رقم (2-8) جودة عنصر الوصف في مقالات الدوريات الإلكترونية العربية في المجال

الدورية	عدد المقالات	تعبر القيم في عنصر الوصف عن المقالة	النسبة	لا تعبر القيم في عنصر الوصف عن المقالة	النسبة
Cybrarians Journal	108	صفر	صفر%	108	%100
العربية 3000	51	صفر	صفر%	51	%100
دراسات المعلومات	42	صفر	صفر%	42	%100
المجلة العربية للدراسات المعلوماتية	15	صفر	صفر%	15	%100
مجلة المكتبات والمعلومات	6	6	%100	صفر	صفر%
الإجمالي	222	6	%2.70	216	%97.30

(1) Krishnan, Raj. Improve snippets with a meta description makeover. [online]. Google Webmaster Central Blog, 2007. Available at: https://webmasters. googleblog. com/2007/09/improve-snippets-with-meta-description.html . [Accessed 24/11/2015].

يتبين من البيانات الواردة في الجدول رقم (2-8) أن:

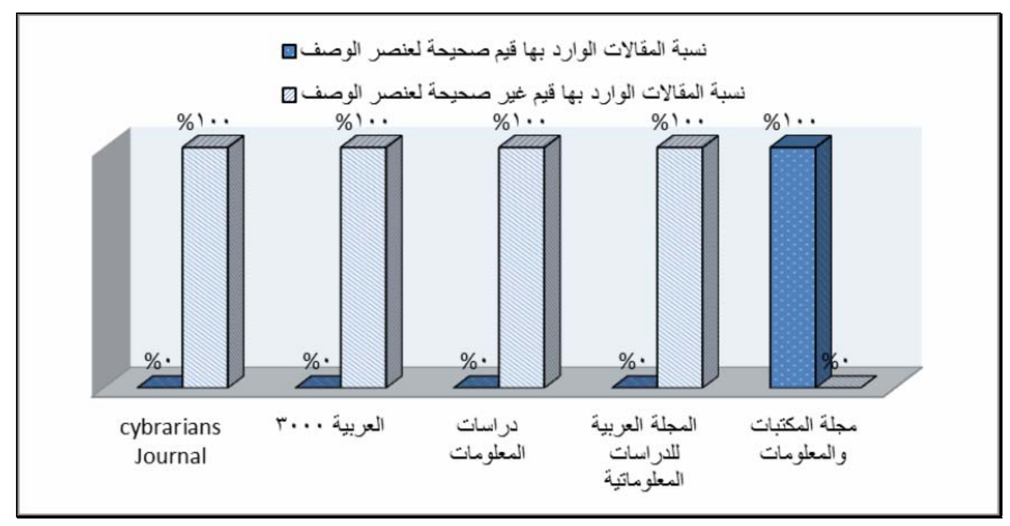
- ورد عنصر الوصف بشكل غير صحيح في معظم مقالات الدوريات الإلكترونية العربية المتخصصة في المكتبات والمعلومات، ولا تعبر القيم الواردة فيه عن وصف هذه المقالات ؛ فقد بلغ عدد المقالات الوارد بها عنصر الوصف بشكل لا يعبر عن كل مقالة (216) مقالة، وهو ما يعادل نحو 97.30% من إجمالي المقالات. بينما كان عدد المقالات الوارد بها عنصر الوصف بشكل مناسب لكل مقالة ست (6) مقالات بنسبة بلغت 2.70%.
- انتمت جميع المقالات الوارد بها عنصر وصف صحيح يتناسب مع كل مقالة- وهي 6 مقالات - إلى مجلة المكتبات والمعلومات، كما أن هذه الدورية تحرص على تسجيل عنصر الوصف بشكل صحيح في صيغة تيجان الميتا وكذلك باستخدام بؤرة دبلن.
- ظهر عنصر الوصف بشكل لا يتناسب مع الوصف الصحيح لكل مقالة في جميع مقالات كل من الدوريات التالية:
  - دراسات المعلومات والمجلة العربية للدراسات المعلوماتية حيث ورد هذا العنصر فارغا في كل منهما، وفي هذا إشارة إلى أن البرنامج المستخدم في إعداد الموقع الذي تتاح عليه الدوريتين - سبقت الإشارة إلى أن الدوريتان تتاحان على موقع واحد - يقوم آليا بإضافة عنصر الوصف ولكن ينبغي على منشئ أو مصمم الموقع أن يستغل هذه الإمكانية ويقوم بتسجيل قيمة مناسبة لهذا العنصر.
  - ورد عنصر الوصف في جميع مقالات دورية Cybrarians Journal بشكل لا يتناسب مع كل مقالة على حدة، وإنما ورد فيه وصف يخص الدورية ككل، كما تجدر الإشارة إلى أن هذا الوصف ورد باللغة الإنجليزية كالتالي:

<meta name="description" content="Cybrarians Journal is refereed e-journal published in Egypt by The Arabic Portal for Librarianship and Information since 2004> ".

أما دورية العربية 3000 فورد عنصر الوصف في جميع مقالاتها وبه كلمة واحدة باللغة الإنجليزية هي "arabcin"، ومن المرجح أن هذه الكلمة هي اختصار للنادي

العربي للمعلومات وهو الاسم السابق لنادي الإحياء العربي الذي تتاح علي موقعه هـذه الدورية، وبالتالي يمكن القول بأن هذا وصف للموقع ككل، وحتى إذا افترضنا ذلك، فإن هذا الوصف للموقع نفسه غير واضح على الإطلاق.

- يتضح من العرض السابق والشكل رقم (2-8) أن مقالات دورية واحدة فقـط هـي مجلة المكتبات والمعلومات العربية ورد بها عنصر الوصف بشكل صحيح في مقابـل أربع دوريات هي Cybrarians Journal، ودراسات المعلومـات، والمجلـة العربيـة للدراسات المعلوماتية، والعربية 3000 ورد بها عنصر الوصف بشكل خاطئ.



الشكل رقم (2-8) جودة عنصر الوصف في مقالات الدوريات الإلكترونية العربية في المكتبات والمعلومات

3-5-2 عنصر الكلمات المفتاحية:

يسجل في هذا العنصر الكلمات المفتاحية التي تعبر عن محتوى الصفحة مفصولة عـن بعضها بفاصلة comma، إلا أن هذا العنصر بعد فترة من ظهوره في ديسمبر من عام 1995، وبسبب بعض الممارسات السيئة التي تمثلت في تكرار كتابة نفس الكلمات المفتاحيـة مـرات عديدة في محاولـة للحصـول عـلى ترتيـب متقـدم في قائمـة نتـائج البحـث، أعلنـت العديـد

من محركات البحث عن التخلي عن هذا العنصر[1]. ومن بين محركات البحث التي لا تدعم هذا العنصر محرك البحث جوجل ؛ وإن كان هناك عدد من محركات البحث التي تدعمه مثل ياهو وآسك[2]. وفيما يلي يوضح الجدول رقم (2-9) جودة عنصر الكلمات المفتاحية لمقالات الدوريات الإلكترونية العربية في المجال.

الجدول رقم (2-9) جودة عنصر الكلمات المفتاحية في مقالات الدوريات الإلكترونية العربية المتخصصة في المجال

الدورية	عدد المقالات	تعبر القيم في عنصر الكلمات المفتاحية عن المقالة	النسبة	لا تعبر القيم في عنصر الكلمات المفتاحية عن المقالة	النسبة
Cybrarians Journal	108	3	%2.78	105	%97.22
العربية 3000	51	صفر	صفر%	51	%100
دراسات المعلومات	42	صفر	صفر%	42	%100
المجلة العربية للدراسات المعلوماتية	15	صفر	صفر%	15	%100
مجلة المكتبات والمعلومات	6	6	%100	صفر	صفر%
الإجمالي	222	9	%4.05	213	%95.96

يتبين من الجدول رقم (2-9) ضعف عنصر الكلمات المفتاحية في مقالات الدوريات الإلكترونية العربية في المجال، فعلى الرغم من أن هذا العنصر ورد في جميع المقالات إلا أن القيم الواردة في النسبة الأكبر منها غير صحيحة ولا تتناسب مع كل مقالة ويمكن تفصيل ذلك كالتالي:

(1) Sullivan, D..Meta Keywords Tag 101: How To "Legally" Hide Words On Your Pages For Search Engines. [online] Search Engine Land, 2007. Available at: http://searchengineland.com/meta-keywords-tag-101-how-to-legally-hide-words-on-your-pages-for-search-engines-12099 [12/2/2016].

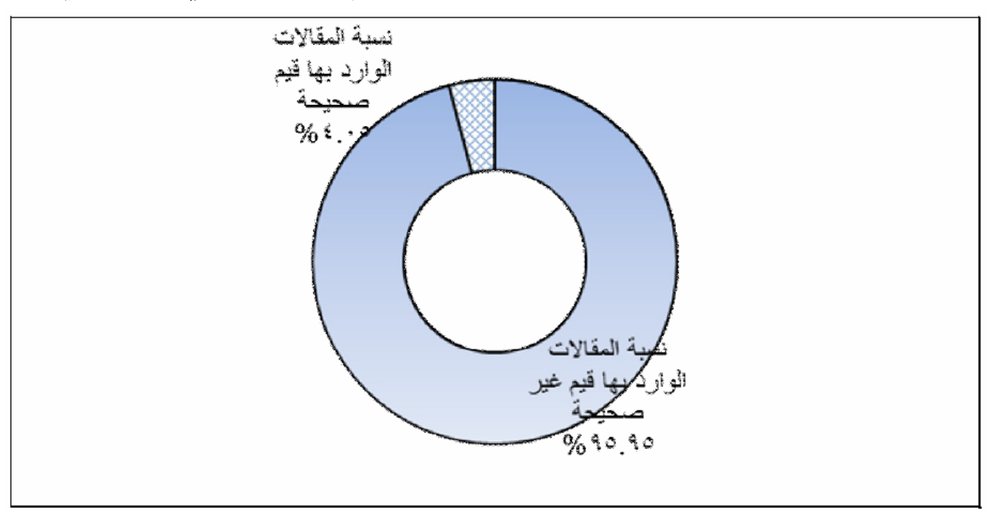
(2) Ibid.

- بلغ عدد المقالات الإلكترونية التي ورد بها كلمات مفتاحية لا علاقة لها بها (213) مقالة بنسبة حوالي 95.95 %، بينما كان عدد المقالات الإلكترونية التي ورد بها كلمات مفتاحية تعبر عنها (9) مقالات أي ما يساوي 4.05% من إجمالي المقالات محل الدراسة.
- ورد عنصر الكلمات المفتاحية بشكل غير صحيح في مقالات كل من دوريتي دراسات المعلومات والمجلة العربية للدراسات المعلوماتية حيث جاء هذا العنصر فارغا، كما أن هذا العنصر في العربية 3000 احتوى على كلمة واحدة باللغة الانجليزية هي "arabcin" وهي نفس الكلمة التي سبق استخدامها في عنصر الوصف. أما دورية Cybrarians Journal فقد ورد في الجزء الأكبر من مقالاتها (97.22%) كلمات مفتاحية تعبر عن الدورية نفسها ولا تعبر عن كل مقالة على حدة، حيث سجلت خمس عشرة (15) كلمة مفتاحية جميعها باللغة الإنجليزية كالتالي:

<meta name="keywords" content="cybrarians, e-journal, electronic, researches, papers, library, information, science, Egypt, librarians, technology, digital, web 2.0, education, cataloging">

- ورد عنصر الكلمات المفتاحية بشكل صحيح في جميع مقالات مجلة المكتبات والمعلومات والتي حرصت على إدراج هذا العنصر باستخدام تيجان الميتا، كما ورد باستخدام عنصر بؤرة دبلن الموضوع Subject، ويبدو أن الدورية تنتهج منهجا ثابتا إزاء الكلمات المفتاحية التي تستخدمها حيث لوحظ استخدام خمس كلمات مفتاحية لوصف كل مقالة ترد جميعها باللغة العربية.
- ورد عنصر الكلمات المفتاحية بصورة صحيحة وبشكل استثنائي في ثلاث مقالات فقط من إجمالي مقالات Cybrarians Journal بنسبة تقدر بحوالي 2.78 % من إجمالي مقالات هذه الدورية مع ملاحظة أن جميع هذه المقالات تنتمي إلى عدد واحد هو العدد الثالث، كما تجدر الإشارة إلى أن عنصر الكلمات المفتاحية ورد في هذه المقالات باللغتين العربية والإنجليزية كما يلي:

<meta name="keywords" content="Bibliographic instruction, التعلـــيم البببليوجرافي" />
<meta name="title" content="برامج التعليم البببليوجرافي في المكتبات الأكاديمية ودورها في دعم البحث العلمي / وليد غالي نصر" />



الشكل رقم (2-9) نسبة المقالات الوارد بها عنصر الكلمات المفتاحية بشكل صحيح مقابل المقالات الوارد بها عنصر الكلمات المفتاحية بشكل غير صحيح

4-5-2 عنصر المؤلف:

ورد عنصر المؤلف في مقالات اثنين فقط من الدوريات الإلكترونية العربية في مجال المكتبات والمعلومات وهما Cybrarians Journal، ومجلة المكتبات والمعلومات. ولذلك فإن النسبة تم حسابها من إجمالي مقالات كل من الدوريتين فقط. مع العلم بأن هذا العنصر يسجل به مؤلف الصفحة، ونظرا لأن لكل مقالة مؤلف خاص بها، فإنه من المفترض أن يتم تسجيله في هذا العنصر.

الجدول رقم (2-10) جودة عنصر المؤلف في مقالات الدوريات الإلكترونية العربية في المجال

الدورية	عدد المقالات	تعبر القيم في عنصر المؤلف عن المقالة	النسبة	لا تعبر القيم في عنصر الوصف عن المقالة	النسبة
Cybrarians Journal	108	صفر	صفر%	108	100%
مجلة المكتبات والمعلومات	6	6	100%	صفر	صفر%
الإجمالي	114	6	5.26%	108	94.74%

- ظهرت قيم تاج المؤلف في جميع مقالات Cybrarians Journal بطريقة لا تتناسب مع هذه المقالات، حيث تم فيه تسجيل اسم الشخص المسؤول عن تحرير ونشر المقالة على موقع الدورية، والذي عادة ما يكون له حساب باسم مستخدم وكلمة سر يستطيع من خلالها تسجيل الدخول على برنامج إدارة المحتوى جوملا Joomla – وهو البرنامج المستخدم في إنشاء موقع الدورية كما سبق الإشارة إلى ذلك – وبالتالي يمكن استنتاج أن قيم هذا العنصر تسجل بياناته آليا وفقا لصاحب الحساب الذي يقوم بنشر أو إضافة المقالة على موقع الدورية، ولذلك فكانت القيم الواردة في هذا العنصر في جميع مقالات هذه الدورية موزعة كالتالي: Administrator (ورد في 23 مقالة)، وAziza Aly (ورد في 34 مقالة)، وMahmoud Khalifa (ورد في 51 مقالة).
- حرصت مجلة المكتبات والمعلومات على تسجيل عنصر المؤلف بشكل صحيح يتناسب مع كل مقالة على حدة، حيث تم تسجيل مؤلف كل مقالة سواء باستخدام عنصر المؤلف author لتيجان الميتا أو باستخدام عنصر المنشئ creator لبؤرة دبلن.

2-5-5 عنصر التاريخ واللغة والمحدد والمصدر:

انفردت دورية واحدة فقط من الدوريات محل الدراسة وهي مجلة المكتبات والمعلومات بتسجيل كل من عنصر التاريخ واللغة والمحدد والمصدر في جميع مقالاتها، والتي شكلت إجمالي المقالات الوارد بها هذا العنصر من إجمالي المقالات محل الدراسة. ومن بين هذه

العناصر ورد عنصر اللغة فقط باستخدام تيجان الميتا HTML meta tags، وباستخدام بؤرة دبلن. فيما وردت باقي العناصر التاريخ والمحدد والمصدر بصيغة بؤرة دبلن فقط.

الجدول رقم (2-11) جودة عنصر التاريخ واللغة والمحدد والمصدر في مقالات مجلة المكتبات والمعلومات

عنصر ما وراء البيانات	تعبر القيم في هذا العنصر عن المقالة	النسبة	لا تعبر القيم في هذا العنصر عن المقالة	النسبة
التاريخ	6	100%	صفر	صفر%
اللغة	صفر	صفر%	6	100%
المحدد	6	100%	صفر	صفر%
المصدر	6	100%	صفر	صفر%

يتبين من الجدول رقم (2-11) ما يلي:

- جاء عنصر التاريخ بصيغة بؤرة دبلن في جميع مقالات هذه الدورية، وجاءت قيم هذا العنصر مناسبة لجميع مقالات هذه الدورية أي بنسبة 100%، إلا أنه تجدر الإشارة إلى أن التاريخ الوارد في هذا العنصر يشير إلى تاريخ نشر المقالة على الموقع وليس تاريخ نشر المقالة في الإصدارة الورقية، فهذه الدورية تصدر في الأصل في شكل ورقي، وبالتالي فإن المقالة المتاحة من خلال هذه الدورية في صيغة بي دي إف PDF تحمل في الأصل تاريخ نشر العدد الورقي، وقد تمت صياغة التاريخ وفقا لبنية الأيزو 8601 ISO8601 Schema.
- لم تعبر القيمة في عنصر اللغة عن المقالات في هذه الدورية بطريقة صحيحة، سواء ورد بصيغة تيجان الميتا أو بصيغة بؤرة دبلن. فقد بلغت نسبة المقالات الوارد بها هذا العنصر بشكل غير صحيح نسبة 100% ؛ حيث كانت القيمة المسجلة في هذا العنصر لجميع مقالات هذه الدورية هي رمز اللغة الفرنسية fr، وهي القيمة التي تتناسب مع لغة موقع بوابة الدوريات التي تتاح من خلالها هذه الدورية، والتي لا تتناسب مع

لغة المقالات وهي اللغة العربية. ويمكن استنتاج أن هـذه القيمـة قـد تـم إدراجهـا آليـا باستخدام برنامج إدارة المحتوى للموقع المتاحة من خلاله هذه الدورية.

- عادة ما يسجل في عنصر المحـدد الموقـع الموحـد للمصـدر URL، وقـد وردت قـيم هـذا العنصر صحيحة في جميع مقالات مجلة المكتبات والمعلومات أي ما نسبته 100%، حيث ورد في هذا العنصر المحدد الخاص بكل مقالة.
- بلغ عدد المقالات التي ورد بها قيم تتناسب مـع كـل مقالـة في عـنصر المصـدر سـت (6) مقالات بما يعادل 100% من المقالات التي ورد بها هذا العنصر، إلا إن استخدام العـنصر نفسه لا يعد صحيحا نسبيا، فهذا العنصر استخدمته مجلة المكتبات والمعلومات لتسجيل موقعها باعتبارها مصدرا لهذه المقالات، حيث ورد هذا العنصر كالتالي:

<meta name="DC.source" scheme="DCTERMS.URI"
content="http://www.webreview.dz">

إلا أن بؤرة دبلن أشارت إلى أن عنصر المصدر source يستخدم في وصف العلاقة بين المصدر الموصوف ومصدر آخر تم منه اشتقاق المصـدر الموصـوف بشـكل كامـل أو جـزئي، على سبيل المثال كتاب ما أو دورية تعد مصدرا لنسخة ضوئية رقمية[1]. أي بعبـارة أخـرى أن يذكر في عنصر المصدر الأصل الذي تم منه اشتقاق المصدر الموصوف بيـنما أشـار إلى أن عنصر جزء من IsPartOf يستخدم لوصف العلاقة بين المصدر الموصوف (وهو في الدراسة الحالية المقالات) ومصدر آخر يكون فيـه المصـدر الموصـوف جـزء مـادي أو منطقـي منـه (على سبيل المثال لوحة كجزء من مجموعة، مقالة كجزء من دورية، إلخ) [2]. وبالتالي يكـون من الأنسب استخدام عـنصر جـزء مـن IsPartOf لوصـف بيانـات الدوريـة التـي تـرد بهـا المقالة. وقد يرجع استخدام مجلة المكتبات والمعلومات لعـنصر المصـدر بهـذه الطريقـة إلى

(1) Rühle, Stefanie, Baker, Tom, Johnston, Pete. User Guide/Creating Metadata - DCMI_MediaWiki. [online] Available at: http://wiki.dublincore.org/index.php/User_Guide/Creating_Metadata .[Accessed 18/8/2015].

(2) Ibid.

عدم إلقاء بؤرة دبلن لمزيد من التوضيح حول استخدام هذا العنصر مما يؤدي إلى حدوث لبس حول كيفية استخدامه. وهذا يتفق مع ما أشار إليه بارك Park [1]من أن تعريفات بعض عناصر بؤرة دبلن غامضة ومتداخلة مع بعضها، وبالتحديد عناصر مثل المصدر Source والعلاقة Relation، والشكل Format والنوع type.

3 نظرة استكشافية لعناصر ما وراء البيانات في عدد من مقالات الدوريات الإلكترونية الأجنبية في مجال المكتبات والمعلومات:

يحاول هذا الجزء استكشاف عناصر ما وراء البيانات في عدد من مقالات الدوريات الإلكترونية الأجنبية وذلك للتعرف على توجه هذه الدوريات حيال عناصر ما وراء البيانات بها.

ولذلك تم اختيار ثلاث دوريات إلكترونية أجنبية في مجال المكتبات والمعلومات لاستكشاف عناصر ما وراء البيانات بها وهي: D - Lib Magazine، وAriadne، و First Monday. وهذه الدوريات تم اختيارها بناء على أنها أكثر الدوريات استخداما من قبل الباحثين العرب في مجال المكتبات والمعلومات[2]، كما أن جميع مقالاتها تتاح مجانا وتتوافر بلغة ترميز النص الفائق HTML. ولاستكشاف عناصر ما وراء البيانات في مقالات هذه الدوريات، تم اختيار عينة عشوائية تمثل خمس مقالات من كل دورية بشكل راجع من آخر خمسة أعوام تبدأ من عام 2015 (أي من عام 2015 حتى عام 2011)، وتم اختيار هذه المقالات بواقع مقالة من العدد الأول لكل عام على أن تكون هذه المقالة هي أول مقالة في العدد ترد بعد الافتتاحية إن وجدت. ويوضح الجدول التالي بيانات هذه الدوريات:

---

(1) Park, Jung-ran, Childress, Eric. Dublin Core metadata semantics: an analysis of the perspectives of information professionals. Journal of Information Science. Vol.35, No. 6 (2009). Pp. 727–739.

(2) لطفي، منيرة محمد مظهر. أنماط إفادة الباحثين العرب فى مجال المكتبات والمعلومات من المصادر الإلكترونية للمعلومات / إشراف حشمت قاسم. أطروحة (ماجستير). جامعة القاهرة – كلية الآداب – قسم المكتبات والوثائق والمعلومات، 2009. ص 151.

الجدول رقم (2-12) الدوريات الإلكترونية الأجنبية التي خضعت لاستكشاف عناصر ما وراء البيانات بها

الدورية	D - Lib Magazine	Ariadne	First Monday
تاريخ صدور العدد الأول	يوليو 1995	يناير 1996	مايو 1996
هيئة الإصدار	مؤسسة مبادرات الأبحاث الوطنية – الولايات المتحدة الأمريكية Corporation for National Research Initiatives	مكتبة جامعة لوفبورو بالمملكة المتحدة Loughborough University Library	مكتبة جامعة الينوي في شيكاغو University of Illinois at Chicago - University Library
تتابع الصدور	ست مرات في العام	غير محدد	شهرية
إتاحة النص الكامل للمقالات	مجانا	مجانا	مجانا
الأشكال المتاحة	إلكتروني	إلكتروني	إلكتروني
صيغ عرض المقالات	لغة ترميز النص الفائق HTML	لغة ترميز النص الفائق HTML	لغة ترميز النص الفائق HTML
إصدارة لغة ترميز الدورية	HTML 5	XHTML 1.0 Transitional	XHTML 1.0 Transitional
الموقع الإلكتروني	http://www.dlib.org/dlib.html	http://www.ariadne.ac.uk/	http://firstmonday.org/ojs/index.php/fm

يتبين من بيانات الجدول رقم (2-12) العناصر التالية:

- يتقارب تاريخ إصدار هذه الدوريات ما بين عام 1995 وعام 1996، وهـذا يسـبق تـاريخ صدور أول دورية عربية من الدوريات محل الدراسة وهي العربية 3000 التي صدر أول عدد منها عام 2000.كما أن جميع هذه الدوريات مستمرة في الصدور.
- جميع هذه الدوريات هي دوريات إلكترونية فقط ولا يصدر منها إصدار ورقـي، كـما أن جميع مقالاتها يتاح نصها الكامل مجانا.

1-3 تحليل عناصر ما وراء البيانات في الدوريات الإلكترونية الأجنبية المختارة:

تم تحليل عناصر مـا وراء البيانـات لكـل دوريـة مـن الـدوريات الثلاثـة بواقـع خمـس مقالات، وتجدر الإشارة إلى أنه تم تجاهل عناصر ما وراء البيانات التي وردت كملف مسـتقل يتم ربطه بصفحة العنكبوتية المتاح عليها المقالة، والذي عادة ما يضم عناصر ما وراء بيانـات بلغة الترميز الموسعة XML، وذلـك ورد في دوريـة D – Lib Magazine، حيـث وردت فيهـا بجانب عناصر ما وراء البيانات الواردة في شفرة مصدر الصفحة، ملفات مستقلة تحتوي عـلى عناصر ما وراء بيانات لوصف المقالة ويتم ربطها باستخدام عنصر link كما يلي:

```
<link rel="metadata" href="01brase.meta.xml" />
```

وبالضغط على رابط هذا الملف الموجود في الخاصية href تظهر عناصر ما وراء البيانات كاملة ومناسبة عن المقالة كما في الشكل رقـم (2-10) ويظهـر في هـذه الملفـات المسـتقلة أن هذه الدورية تستخدم خطة ما وراء بيانات خاصة بها تم الإعلان عنهـا في بدايـة هـذا الملـف بـ" dlib-meta0.1"، وبالتالي فهذا الملف قاصر فقط على التطبيقات والـبرامج التـي تسـتطيع تفسيره.

```
▼<dlib-meta0.1>
   <title>Access to Research Data</title>
   <creator>Jan Brase</creator>
   <creator>Adam Farquhar</creator>
   <publisher>Corporation for National Research Initiatives</publisher>
   <date date-type="publication">January/February 2011</date>
   <type resource-type="work">article</type>
   <identifier uri-type="DOI">10.1045/january2011-braae</identifier>
 ▼<identifier uri-type="URL">
     http://www.dlib.org/dlib/january11/brase/01brase.html
   </identifier>
   <language>English</language>
 ▼<relation rel-type="InSerial">
     <serial-name>D-Lib Magazine</serial-name>
     <issn>1082-9873</issn>
     <volume>17</volume>
     <issue>1/2</issue>
   </relation>
   <rights>Copyright 2011 Jan Brase and Adam Farquhar</rights>
 </dlib-meta0.1>
```


الشكل رقم (2-10) ملف مستقل لما وراء البيانات بلغة الترميز الموسعة XML لأحد مقالات دورية D-Lib Magazine

وفيما يلي يوضح الجدول رقم (2-13) مدى توافر عناصر ما وراء البيانات في مقالات الدوريات الإلكترونية الأجنبية في مجال المكتبات والمعلومات

الجدول رقم (2-13) مدى توافر ما وراء البيانات في مقالات الدوريات الإلكترونية الأجنبية في المجال

الدورية	توجد عناصر ما وراء البيانات		لا توجد عناصر ما وراء البيانات	
	العدد	النسبة من مقالات كل دورية	العدد	النسبة من مقالات كل دورية
D - Lib Magazine	5	%100	صفر	صفر%
Ariadne	5	%100	صفر	صفر%
First Monday	5	%100	صفر	صفر%
الإجمالي	15	%100	صفر	صفر%

ظهرت في جميع مقالات الدوريات الأجنبية الثلاثة المتخصصة في مجال المكتبات والمعلومات عناصر ما وراء البيانات، حيث ظهرت عناصر ما وراء بيانات أخرى بخلاف عنصر العنوان الذي يسجل آليا بواسطة برامج تصميم المواقع أو برامج إدارة محتوى المواقع.

وفيما يتعلق بخطط ما وراء البيانات المستخدمة في تسجيل عناصر ما وراء البيانات في مقالات هذه الدوريات، يتبين من الجدول رقم (2-14) أن الخطة الأساسية المستخدمة في إدراج عناصر ما وراء البيانات هي تيجان الميتا HTML Meta tags والتي وردت في مقالات الدوريات الثلاثة أي بنسبة 100% من إجمالي مقالات هذه الدوريات.

الجدول رقم (2-14) خطط ما وراء البيانات في مقالات الدوريات الإلكترونية الأجنبية في مجال المكتبات والمعلومات

خطط ما وراء البيانات / الدورية	عناصر ما وراء بيانات لغة ترميز النص الفائق HTML Metadata		بؤرة دبلن		عناصر ما وراء البيانات لجوجل الباحث العلمي	
	العدد	النسبة من مقالات كل دورية	العدد	النسبة من مقالات كل دورية	العدد	النسبة من المقالات
D - Lib Magazine	5	%100	صفر	صفر%	صفر	صفر%
Ariadne	5	%100	صفر	صفر%	5	%100
First Monday	5	%100	5	%100	5	%100
الإجمالي	15	%100	5	%33.33	10	%66.67

- في المرتبة الثانية من حيث خطط ما وراء البيانات المستخدمة جاءت عناصر ما وراء البيانات التي يُوصي بها جوجل الباحث العلمي Google Scholar وتحديدا تلك التي حددها بتيجان أو عناصر هاي ويربرس Highwire Press tags[1]. حيث بلغ عدد

(1) هاي وير بريس هي شركة أمريكية تم تأسيسها بواسطة مكتبات جامعة ستنافورد Stanford

==

المقالات التي استخدمت هذه الخطة (10) مقالات بنسبة حوالي 66.67% من إجمالي المقالات وذلك في مقالات كل من دورية Ariadne، ودورية First Monday. وقد أوصى محرك البحث جوجل الباحث العلمي Google Scholar بأن ترد هذه العناصر كالتالي :[1]

```
<meta name="citation_title" content="The testis isoform of the phosphorylase kinase
    catalytic subunit (PhK-T) plays a critical role in regulation of glycogen
    mobilization in developing lung">
<meta name="citation_author" content="Liu, Li">
<meta name="citation_author" content="Rannels, Stephen R.">
<meta name="citation_author" content="Falconieri, Mary">
<meta name="citation_author" content="Phillips, Karen S.">
<meta name="citation_author" content="Wolpert, Ellen B.">
<meta name="citation_author" content="Weaver, Timothy E.">
<meta name="citation_publication_date" content="1996/05/17">
<meta name="citation_journal_title" content="Journal of Biological Chemistry">
<meta name="citation_volume" content="271">
<meta name="citation_issue" content="20">
<meta name="citation_firstpage" content="11761">
<meta name="citation_lastpage" content="11766">
<meta name="citation_pdf_url"
    content="http://www.example.com/content/271/20/11761.full.pdf">
```

حيث يعد جوجل الباحث العلمي Google Scholar محرك بحث خاص ببحث

==

University تقدم خدمات لتطوير المحتوى الرقمي وتعمل على توفير خدمات استضافة وحلول لعدد ضخم من المطابع الجامعية والناشرين الذين يقومون بإصدار الدوريات والكتب وغيرها من المنشورات العلمية ؛ حيث تعمل على تقديم استضافة على الانترنت لأكثر من 3500 دورية محكمة، وآلاف من الكتب العلمية. (المصدر:

HighWire Press [online]. In: Wikipedia: The Free Encyclopedia. Available at: https://en.wikipedia.org/wiki/HighWire_Press . [Accessed 17/5/2016].

(1) Google Scholar. Inclusion Guidelines for Webmasters. Indexing Guidelines [online]. Available at: https://scholar.google.com/intl/en/scholar/inclusion.html#indexing . [Accessed 18/5/2016].

الإنتاج العلمي فقط ويشمل البحث في مدى واسع من المصادر العلمية مثل المقالات، والأطروحات، والكتب[1]. كما أن هذا المحرك يختلف في طريقة عرضه لنتائج البحث حيث تصاحب كل نتيجة بيانات ببليوجرافية مثل اسم المؤلف، وتاريخ النشر، وعدد مرات الاستشهاد كما يظهر في الشكل رقم (2-11).

Google
Emergency-relief coordination on social media
الباحث العلمي
حوالي 14,900 من النتائج (0.04 من الثواني)
المقالات
مكتبتي
أي وقت
منذ 2016
منذ 2015
منذ 2012
نصيحة: بحث عن النتائج باللغة **العربية** فقط. يمكنك تحديد لغة البحث في إعدادات الباحث العلمي.
Emergency-relief coordination on **social media**: Automatically matching resource requests and offers
[HTML] من firstmonday.org
firstmonday.org - H Purohit, C Castillo, F Diaz, A Sheth, P Meier - First Monday, 2013
Abstract Disaster affected communities are increasingly turning to social media for communication and coordination. This includes reports on needs (demands) and offers (supplies) of resources required during emergency situations. Identifying and matching ...
تم اقتباسها في عدد: 17 مقالات ذات صلة الإصدارات لـ 18كلها اقتبس حفظ المزيد

الشكل رقم (2-11) نتائج البحث كما تظهر في محرك البحث جوجل الباحث العلمي Google Scholar

وتجدر الإشارة إلى أن هذا المحرك قد أشار أيضا إلى دعمه لبؤرة دبلن، إلا أنه يرى أن هذه الخطة قد تكون قاصرة نظرا لعدم وجود حقول خاصة لوصف عنوان الدورية، والمجلد، والعدد، وأرقام الصفحات[2].

---

(1)Google. About Google Scholar [online].Available at: https://scholar.google.com/intl/en/scholar/about.html .[Accessed 17/5/2016].

(2) Google Scholar. Op. Cit.

ويلاحظ أن مقالات الدوريات الإلكترونية العربية المتخصصة في المجال قد استخدمت كلا من عناصر تيجان الميتا، وعناصر بؤرة دبلن إلا أنه لا يوجد من بينها ما استخدم عناصر ما وراء البيانات الموصى بها من قبل جوجل الباحث العلمي Google Scholar.

- من إجمالي مقالات الدوريات الإلكترونية الأجنبية في مجال المكتبات والمعلومات، كان عدد المقالات التي استخدمت بؤرة دبلن لإدراج عناصر ما وراء البيانات بها خمس (5) مقالات فقط، أي ما يوازي نحو 33.33% من إجمالي العينة لمقالات هذه الدوريات، وقد انتمت جميع هذه المقالات لدورية واحدة فقط هي دورية First Monday والتي استخدمت تيجان الميتا أيضا بجانب استخدامها لبؤرة دبلن. كما أن هذه الدورية لم تتوقف عند استخدام العناصر الأساسية لبؤرة دبلن بل استخدمت كذلك محددات أو مقيدات بؤرة دبلن Dublin Core Qualifiers مع عدد من العناصر. فعلى سبيل المثال مع عنصر التاريخ date تم استخدام مقيدات مثل التاريخ. الإنشاء date.created، والتاريخ.الإيداع date.submitted، والتاريخ. الإصدار date.issued، إلا أنه ينبغي الإشارة إلى أن هذه المحددات حلت محلها مفردات بؤرة دبلن DCMI Metadata Terms في عام 2012[(1)]، وبالتالي فالأصح هو استخدام هذه العناصر وفقا لهذه المفردات كالتالي: DCTERMS. created، وDCTERMS. Date Submitted، وDCTERMS. Issued.

2-3 التوزيع النوعي لعناصر (تيجان) ما وراء البيانات في مقالات الدوريات الإلكترونية الأجنبية:

تم حصر أنواع عناصر ما وراء البيانات التي وردت في مقالات الدوريات الإلكترونية الأجنبية الثلاثة، والتي يوضحها الجدول رقم (2-15)، مع ملاحظة أنه تم حساب نوع العنصر بغض النظر عن الخطة الوارد بها، فإذا ما ورد عنصر الوصف وفقا لأكثر من خطة

(1)Dublin Core Qualifiers [online]. 11-07-2000. Available at: http://dublincore.org/documents/2000/07/11/dcmes-qualifiers/ . Accessed [22/10/2013].

الجدول رقم (2-15) التوزيع النوعي لعناصر ما وراء البيانات في مقالات الدوريات الإلكترونية الأجنبية في مجال المكتبات والمعلومات

أنواع عناصر ما وراء البيانات	مقالات D - Lib Magazine (5 مقالات)	مقالات Ariadne (5مقالات)	مقالات First Monday (5 مقالات)	عدد الدوريات التي استخدمت هذا العنصر	النسبة من إجمالي عدد الدوريات (3 دوريات)	إجمالي عدد مرات ظهور العنصر	النسبة من إجمالي المقالات (15)
العنوان title	✓	✓	✓	3	%100	15	%100
الوصف description	✓	✓	✓	3	%100	15	%100
المؤلف / المنشيء author /creator	✗	✓	✓	2	%66.67	10	%66.67
التاريخ date	✗	✓	✓	2	%66.67	10	%66.67
المصدر / source عنوان الدورية citation_journal_title	✗	✓	✓	2	%66.67	10	%66.67
العدد issue	✗	✓	✓	2	%66.67	10	%66.67
الترقيم الدولي الموحد للدورية ISSN	✗	✓	✓	2	%66.67	10	%66.67
الكلمات المفتاحية / الموضوع keywords / subject	✓	✗	✓ (4 مقالات)	2	%66.67	9	%60
المجلد volume	✗	✗	✓	1	%33.33	5	%33.33
اللغة language	✗	✗	✓	1	%33.33	5	%33.33

أنواع عناصر ما وراء البيانات	مقالات D - Lib Magazine (5 مقالات)	مقالات Ariadne (5مقالات)	مقالات First Monday (5 مقالات)	عدد الدوريات التي استخدمت هذا العنصر	النسبة من إجمالي عدد الدوريات (3 دوريات)	إجمالي عدد مرات ظهور العنصر	النسبة من إجمالي المقالات (15)
المحدد identifier	×	×	✓	1	%33.33	5	%33.33
المعرف الرقمي للكيان DOI	✓	×	✓	1	%33.33	5	%33.33
موقـــــــع ملخـــــــص المقالـــــــة citation_abstract_html_url	×	×	✓	1	%33.33	5	%33.33
نوع المصدر type	×	×	✓	1	%33.33	5	%33.33
الحقوق rights	×	×	✓	1	%33.33	5	%33.33
التغطية coverage	×	×	✓ (2 مقالة)	1	%33.33	2	%13.33
المساهم contributor	×	×	(2 مقالة)	1	%33.33	2	%13.33
ب- فئة عناصر ما وراء البيانات الفنية							
نوع المحتوى / الشكل format	✓	✓	✓	3	%100	15	%100
المولد generator	×	✓	×	1	%33.33	5	%33.33
gs_meta_revision	×	×	✓	1	%33.33	5	%33.33

مثل بؤرة دبلن، وتيجان الميتا على سبيل المثال، فتم حساب العنصر مرة واحدة فقط، كما أنه في حالة استخدام الدورية كما هو الحال في دورية First Monday للمحددات مع عناصر بؤرة دبلن تم حساب العنصر الأساسي بغض النظر عن المحددات المصاحبة له.

من الجدول السابق رقم (2-15) يتبين التالي:

- بلغ إجمالي عدد عناصر ما وراء البيانات الواردة في مقالات الدوريات الإلكترونية الأجنبية (20) عنصرا، توزعت على فئتين هما: أ - عناصر ما وراء البيانات الوصفية وتشمل (17) عنصرا هي (العنوان title، والوصف description، والكلمات المفتاحية / الموضوع keywords/subject، والمؤلف author/creator، والتاريخ date، والمصدر source، والعدد issue، والمجلد volume، والترقيم الدولي الموحد للدورية ISSN، واللغة language، والمحدد identifier، والمعرف الرقمي للكيان DOI، وموقع ملخص المقالة citation_abstract_html_url، ونوع المصدر type، والحقوق rights، والتغطية coverage، والمساهم contributor)، وب – عناصر ما وراء البيانات الفنية وتشمل (3) عناصر هي (نوع المحتوى / الشكل content type/format، والمولد generator، ورقم إصدارة عناصر ما وراء البيانات لجوجل الباحث العلمي gs_meta_revision). وهذا العدد من العناصر يزيد عن الوارد في مقالات الدوريات الإلكترونية العربية ؛ والتي ورد بها (13) عنصرا.
- اشتركت كل من مقالات الدوريات الأجنبية والعربية المتخصصة في المجال في عدد من العناصر بلغت عشرة (10) عناصر شملت: العنوان، والوصف، والكلمات المفتاحية/ الموضوع، والمؤلف/المنشئ، والتاريخ، واللغة، والمحدد، والمصدر، ونوع المحتوى، والمولد.
- في فئة عناصر ما وراء البيانات الوصفية، انفردت مقالات الدوريات الإلكترونية الأجنبية بتسعة عناصر لم ترد في مقالات الدوريات الإلكترونية العربية شملت: بيانات العدد، وبيانات المجلد، والترقيم الدولي الموحد للدورية، والمعرف الرقمي للكيان، وموقع ملخص المقالة، ونوع المصدر، والحقوق، والتغطية، والمساهم. ولم

تشر الدراسة الحالية إلى الفرق بين عناصر ما وراء البيانات في فئة العناصر الفنيـة في كـل من مقالات الدوريات الأجنبية والعربية نظرا لأن هذه العناصر عادة مـا تسـجل بشـكل آلي وفقـا لبرنـامج تصـميم الموقـع، وكـذلك بمـا يتناسـب مـع طبيعـة موقـع كـل دوريـة واحتياجاته.

- كانت أكثر عناصر ما وراء البيانـات ورودا في مقـالات الـدوريات الإلكترونيـة الأجنبيـة في مجال المكتبات والمعلومات، والتي احتلت المرتبة الأولى كل من: العنوان، والوصف، ونوع المحتوى والتي جاءت في جميع مقالات هذه الدوريات بنسبة 100%، وهذا يتوافـق مـع كون هذه العناصر أكثر العناصر ورودا في مقالات الـدوريات الإلكترونيـة العربيـة حيـث وردت كذلك في جميع مقالات هذه الدوريات.
- جـاء في المرتبـة الثانيـة وبنسـبة بلغـت نحـو 66.67% مـن مقـالات الـدوريات الأجنبيـة المختارة كل من: عنصر المؤلف، والتاريخ، والمصـدر / عنـوان الدوريـة، والعـدد، والترقـيم الدولي الموحـد للدرويـة؛ حيـث اسـتخدمت كـل مـن دوريـة Ariadne، ودوريـة First Monday هذه العناصر في جميع مقالاتها.
- كانت نسبة ورود كل من عنصر المؤلف، والتاريخ، والمصدر أعـلى مـن نسـبة ورود هـذه العناصر في مقالات الدوريات الإلكترونيـة العربيـة ؛ حيـث ورد عـنصر المؤلـف في نسـبة بلغت 51.35% من مقالات الدوريات الإلكترونية العربية، وورد كل من التاريخ والمصـدر والمحدد في مقالات دورية واحدة هي مجلة المكتبات والمعلومات وبنسبة بلغت 2.70%.
- جاء في المرتبة الثالثة عنصر الكلمات المفتاحية في تسع (9) مقـالات بنسـبة بلغـت 60% وذلك في جميع مقالات دوررية D – Lib Magazine، و أربع مقالات فقط مـن دوريـة First Monday، وهذه النسبة أقل من النسبة التي حققها عنصر الكلمات المفتاحيـة في مقالات الدوريات الإلكترونية العربية حيـث ورد هـذا العـنصر في جميـع مقـالات هـذه الدوريات.
- جاء في المرتبة الرابعة بنسبة 33.33% من إجمالي مقالات الدوريات الأجنبية الثلاثـة كـل

من عنصر: المجلد، واللغة، والمحدد، ومحـدد ملخـص المقالـة، ونـوع المصـدر، والحقـوق، ورقم إصـدارة عنـاصر مـا وراء البيانـات لجوجـل الباحـث العلمـي gs_meta_revision (وردت جميع هذه العناصر في مقالات دورية واحدة فقـط هـي First Monday)، كـما ورد في نفس هذه المرتبة كل من عنصر المعرف أو المحـدد الرقمـي للكيـان DOI والـذي استخدمته دورية D-Lib Magazine في جميع مقالاتها، وكذلك عنصر المولد والذي ظهر في مقالات دورية Ariadne.

- جاء في المرتبة الخامسة والأخيرة كل مـن عـنصري التغطيـة والمسـاهم وذلـك في مقـالتين فقط من دورية First Monday، أي ما يعادل نسبة 13.33% من إجمالي المقالات.
- يتضح من الجدول التالي رقم (2-16) تفوق دورية First Monday من حيث عدد أنواع عناصر ما وراء البيانات المستخدمة في مقالاتها؛ حيث ورد في مقالات هـذه الدوريـة (18 عنصرا) – مع ملاحظة أنه تم حساب عدد هذه العناصر بغض النظر عن الخطة المسجلة بها، وبغض النظر عن المقيدات المصاحبة لهذه العناصر -، تلتها دوريـة Ariadne والتـي ورد بها (9 عناصر)، ووردت أخيرا دورية D – Lib Magazine بخمسة (5) عناصر فقط.

الجدول رقم (2-16) التوزيع العددي لعناصر ما وراء البيانات في مقالات الدوريات الإلكترونية الأجنبية الثلاثة

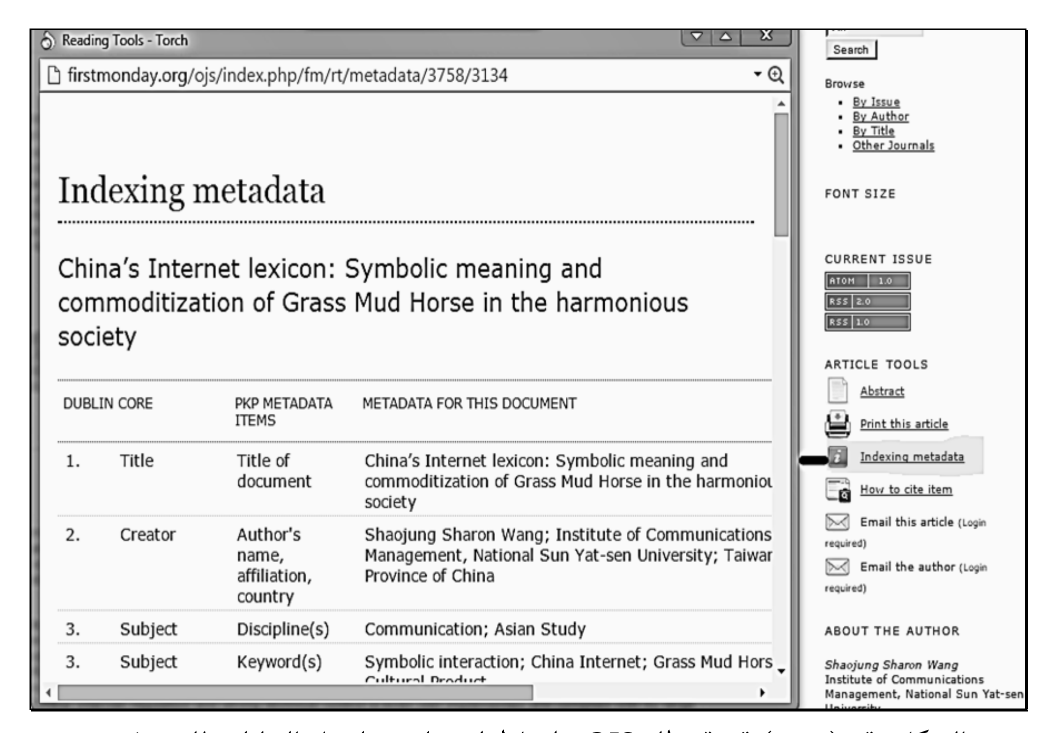
الدورية	عدد أنواع عناصر ما وراء البيانات الواردة في مقالاتها
First Monday	18
Ariadne	9
D – Lib Magazin	5

ومن الجدير بالذكر، أن دراسة سابقة حول دروية First Monday قد أشارت إلى وجود سياسة لما وراء البيانات خاصة بهذه الدورية، وأن عـدد التيجـان أو عنـاصر مـا وراء البابينـات التـي يجـب تسـجيلها لكـل مقالـة هـي تسـعة تيجـان، إلا أن هـذه الدراسـة تقـدم بيانـات

تعد قديمة نسبيا – نشرت المقالة عام 1999- إلا أنها تدل على اهتمام الدورية ومنذ وقت مبكر بما وراء البيانات[1]. وقد يرجع السبب في تفوق هذه الدورية في إدراج عناصر ما وراء البيانات بها إلى أنها تستخدم نظاما خاصا بنشر وإدارة الدوريات الإلكترونية يوفر إمكانات متعددة لإضافة عناصر ما وراء البيانات لمقالات الدوريات، وهذا النظام هو Open Journal Systems (OJS) والذي يعد أحد النظم المفتوحة المصدر والمتاحة مجانا للجميع، وقد تم إنشاء هذا النظام من قبل المشروع العام للمعرفة Public Knowledge Project (PKP) من خلال الجهود التي تمولها الحكومة الفيدرالية في الولايات المتحدة الأمريكية لتوسيع وتحسين فرص الحصول على الأبحاث العلمية، ويهدف هذا النظام إلى جعل النشر مفتوح المصدر خيارا قابلا للتطبيق أمام المزيد من الدوريات[2]، كما يلاحظ أن هذه الدورية بخلاف إدراجها لعناصر ما وراء البيانات في شفرة المصدر Source code لمقالاتها، فإن نظام OJS الذي تستخدمه هذه الدورية يتيح مع كل مقالة رابط بعنوان "indexing metadata" بالضغط عليه يظهر للمستفيد نافذة بها عناصر ما وراء البيانات الكاملة لهذه المقالة – انظر الشكل رقم (2-12).

(1) Henshaw, Robin. The First Monday Metadata Project [online] . Libri. Vol. 49 (1999). Pp.125–131. Available at: http://www.degruyter.com/dg/viewarticle.fullcontentlink:pdfeventlink/$002fj$002flibr.1999.49.issue-3$002flibr.1999.49.3.125$002flibr.1999.49.3.125.pdf/libr.1999.49.3.125.pdf?t:ac=j$002flibr.1999.49.issue-3$002flibr.1999.49.3.125$002flibr.1999.49.3.125.xml. [Accessed 30/5/2016].

(2) Open Journal Systems | Public Knowledge Project. [online] Available at: https://pkp.sfu.ca/ojs/ [Accessed 17/5/2016].



الشكل رقم (2-12) قدرة نظام OJS على إظهار عناصر ما وراء البيانات للمستفيد في دورية First Monday

3-3 جودة ما وراء البيانات في مقالات الدوريات الإلكترونية الأجنبية في مجال المكتبات والمعلومات:

فيما يتعلق بقياس جودة ما وراء البيانات في مقالات الدوريات الإلكترونية الأجنبية الثلاثة في مجال المكتبات والمعلومات، فقد تم التعرض لجودة عناصر ما وراء البيانات بشكل إجمالي وبصفة أساسية العناصر التي تأتي في فئة عناصر ما وراء البيانات الوصفية حيث أن عدد المقالات المدروسة في هذه الدوريات هو عدد قليل - مقارنة بعدد مقالات الدوريات الإلكترونية العربية - ويرجع السبب في ذلك أن الهدف الأساسي كان استكشاف عناصر ما وراء البيانات الواردة في مقالات هذه الدوريات الإلكترونية. ويوضح الجدول رقم (2-17) جودة عناصر ما وراء البيانات الواردة في مقالات هذه الدوريات إجمالا.

الجدول رقم (2-17) جودة ما وراء البيانات في مقالات الدوريات الإلكترونية الأجنبية في مجال المكتبات والمعلومات**

الدورية / أنواع عناصر ما وراء البيانات	مقالات D - Lib Magazine (5 مقالات)		مقالات Ariadne (5 مقالات)		مقالات First Monday (5 مقالات)		إجمالي المقالات	
	تعبر القيم في عنصر ما وراء البيانات عن المقالة	لا تعبر القيم في عنصر ما وراء البيانات عن المقالة	تعبر القيم في عنصر ما وراء البيانات عن المقالة	لا تعبر القيم في عنصر ما وراء البيانات عن المقالة	تعبر القيم في عنصر ما وراء البيانات عن المقالة	لا تعبر القيم في عنصر ما وراء البيانات عن المقالة	تعبر القيم في عنصر ما وراء البيانات عن المقالة	لا تعبر القيم في عنصر ما وراء البيانات عن المقالة
العنوان title	100%	صفر%	100%	صفر%	100%	صفر%	100%	صفر%
الوصف description	صفر	100%	100%	صفر%	100%	صفر%	66.67%	33.33%
الكلمات المفتاحية / الموضوع keywords / subject	40% (2 مقالة)	60% (3 مقالة)	لم يرد هذا العنصر في مقالات هذه الدورية		100% (4 مقالات)	صفر%	66.67%	33.33%
المؤلف / المنشيء author /creator	لم يرد هذا العنصر في مقالات هذه الدورية		100%	صفر%	100%	صفر%	100%	صفر%

** تجدر الإشارة إلى أنه في حالة العناصر التي لم ترد في جميع مقالات الدوريات الإلكترونية الثلاثة، تم حساب النسبة من إجمالي المقالات التي توافر بها هذا العنصر فقط.

الدورية / أنواع عناصر ما وراء البيانات	مقالات D - Lib Magazine (5 مقالات)		مقالات Ariadne (5 مقالات)		مقالات First Monday (5 مقالات)		إجمالي المقالات	
	تعبر القيم في عنصر ما وراء البيانات عن المقالة	لا تعبر القيم في عنصر ما وراء البيانات عن المقالة	تعبر القيم في عنصر ما وراء البيانات عن المقالة	لا تعبر القيم في عنصر ما وراء البيانات عن المقالة	تعبر القيم في عنصر ما وراء البيانات عن المقالة	لا تعبر القيم في عنصر ما وراء البيانات عن المقالة	تعبر القيم في عنصر ما وراء البيانات عن المقالة	لا تعبر القيم في عنصر ما وراء البيانات عن المقالة
التاريخ date	لم يرد هذا العنصر في مقالات هذه الدورية		100%	صفر%	100%	صفر%	100%	صفر%
المصدر source / عنوان الدورية citation_journal_title	لم يرد هذا العنصر في مقالات هذه الدورية		100%	صفر%	100%	صفر%	100%	صفر%
العدد issue	لم يرد هذا العنصر في مقالات هذه الدورية		100%	صفر%	100%	صفر%	100%	صفر%
الترقيم الدولي الموحد للدورية ISSN	لم يرد هذا العنصر في مقالات هذه الدورية		100%	صفر%	100%	صفر%	100%	صفر%
المجلد volume	لم يرد هذا العنصر في مقالات هذه الدورية		لم يرد هذا العنصر في مقالات هذه الدورية		100%	صفر%	100%	صفر%

الدورية / أنواع عناصر ما وراء البيانات	مقالات D - Lib Magazine (5 مقالات)		مقالات Ariadne (5 مقالات)		مقالات First Monday (5 مقالات)		إجمالي المقالات	
	تعبر القيم في عنصر ما وراء البيانات عن المقالة	لا تعبر القيم في عنصر ما وراء البيانات عن المقالة	تعبر القيم في عنصر ما وراء البيانات عن المقالة	لا تعبر القيم في عنصر ما وراء البيانات عن المقالة	تعبر القيم في عنصر ما وراء البيانات عن المقالة	لا تعبر القيم في عنصر ما وراء البيانات عن المقالة	تعبر القيم في عنصر ما وراء البيانات عن المقالة	لا تعبر القيم في عنصر ما وراء البيانات عن المقالة
اللغة language	لم يرد هذا العنصر في مقالات هذه الدورية		لم يرد هذا العنصر في مقالات هذه الدورية		%100	صفر%	%100	صفر%
المحدد identifier	لم يرد هذا العنصر في مقالات هذه الدورية		لم يرد هذا العنصر في مقالات هذه الدورية		%100	صفر%	%100	صفر%
المعرف الرقمي للكيان DOI	%100	صفر	لم يرد هذا العنصر في مقالات هذه الدورية		لم يرد هذا العنصر في مقالات هذه الدورية		%100	صفر%
موقع ملخص المقالة citation_abstract_html_url	لم يرد هذا العنصر في مقالات هذه الدورية		لم يرد هذا العنصر في مقالات هذه الدورية		%100	صفر%	%100	صفر%
نوع المصدر type	لم يرد هذا العنصر في مقالات هذه الدورية		لم يرد هذا العنصر في مقالات هذه الدورية		%100	صفر%	%100	صفر%

الدورية / أنواع عناصر ما وراء البيانات	مقالات D - Lib Magazine (5 مقالات)		مقالات Ariadne (5 مقالات)		مقالات First Monday (5 مقالات)		إجمالي المقالات	
	تعبر القيم في عنصر ما وراء البيانات عن المقالة	لا تعبر القيم في عنصر ما وراء البيانات عن المقالة	تعبر القيم في عنصر ما وراء البيانات عن المقالة	لا تعبر القيم في عنصر ما وراء البيانات عن المقالة	تعبر القيم في عنصر ما وراء البيانات عن المقالة	لا تعبر القيم في عنصر ما وراء البيانات عن المقالة	تعبر القيم في عنصر ما وراء البيانات عن المقالة	لا تعبر القيم في عنصر ما وراء البيانات عن المقالة
الحقوق rights	لم يرد هذا العنصر في مقالات هذه الدورية		لم يرد هذا العنصر في مقالات هذه الدورية		%100	صفر%	%100	صفر%
التغطية coverage	لم يرد هذا العنصر في مقالات هذه الدورية		لم يرد هذا العنصر في مقالات هذه الدورية		%100	صفر%	%100	صفر%
المساهم contributor	لم يرد هذا العنصر في مقالات هذه الدورية		لم يرد هذا العنصر في مقالات هذه الدورية		%100	صفر%	%100	صفر%

ويتبين من الجدول رقم (2-17) عن جودة عناصر ما وراء البيانات في مقالات الدوريات الإلكترونية الأجنبية محل الدراسة أن:

عنصر العنوان:

ورد عنصر العنوان في مقالات الدوريات الثلاثة بجودة بلغت 100%؛ حيث وردت جميع عناوين المقالات مناسبة لكل مقالة، ويلاحظ أن كل من دورية Ariadne، ودورية First Monday عند تسجيل قيمة العنوان وبخاصة في تاج الميتا <title>، يتم تسجيل عنوان الدورية مع عنوان المقالة كالتالي:

<title>Developing Adaptable, Efficient Mobile Library Services: Librarians as Enablers | Ariadne</title>

كما أن دورية First Monday أضافت لعنوان الدورية، اسم العائلة لمؤلف المقالة كالتالي:

<title>Censorship is futile possible but difficult: A study in algorithmic ethnography | Watters | First Monday</title>

ويلاحظ أن هذا هو نفس التوجه الذي قامت به كل من دورية Cybrarians Journal، ومجلة المكتبات والمعلومات - كما ورد سابقا. إلا أنه عند تسجيل قيمة العنوان باستخدام بؤرة دبلن أو عناصر ما وراء البيانات لجوجل الباحث العلمي في مقالات كل من Ariadne، وFirst Monday فقد تم الالتزام بتسجيل العنوان فقط.

عنصر الوصف:

وردت قيم صحيحة في عنصر الوصف في مقالات الدوريات الإلكترونية الأجنبية بنسبة بلغت حوالي 66.67% (10 مقالات) وذلك في مقالات كل من دورية Ariadne، ودورية First Monday، وتجدر الإشارة إلى أن دورية First Monday جاء بها عنصر الوصف مرة وفقا لتيجان الميتا HTML Meta tags، ومرة أخرى وفقا لبؤرة دبلن. ومع استخدام تيجان الميتا كانت قيمة عنصر الوصف غير مناسبة حيث وضع عنوان المقالة، إلا أنه مع عناصر بؤرة دبلن وضع وصفا مناسبا لكل مقالة. وفيما يتعلق بنسبة المقالات التي

وردت بها قيم غير صحيحة لعنصر الوصف، فقد بلغت هذه النسبة حوالي 33.33% (5 مقالات) جاءت كلها في دورية D – Lib Magazine، حيث تم تسجيل عنوان المجلة D – Lib Magazine في هذا العنصر كما يلي:

<meta name="description" content="D-Lib Magazine">

عنصر الكلمات المفتاحية:

ترتفع نسبة المقالات الوارد بها عنصر الكلمات المفتاحية بشكل صحيح عن تلك النسبة التي يرد بها هذا العنصر بشكل غير صحيح؛ حيث بلغت نسبة المقالات التي ورد بها عنصر الكلمات المفتاحية بشكل صحيح بما يتناسب مع كل مقالة نحو 66.67% (6 مقالات) من إجمالي المقالات التي ورد بها هذا العنصر (9 مقالات)، وجاءت هذه المقالات في مقالتين فقط من مقالات دورية D-Lib Magazine، وفي جميع مقالات دورية First Monday والتي حرصت على تسجيل قيم صحيحة لهذا العنصر سواء عند استخدامها لتيجان الميتا، أو لبؤرة دبلن، أو لعناصر ما وراء البيانات لجوجل الباحث العلمي كما يلي:

<title>Censorship is futile possible but difficult: A study in algorithmic ethnography | Watters | First Monday</title>

<meta name="keywords" content="Censorship; Behavioural Advertising; Malware; Risk">

<meta name="DC.Subject" xml:lang="en" content="Censorship">

<meta name="DC.Subject" xml:lang="en" content="Behavioural Advertising">

<meta name="DC.Subject" xml:lang="en" content="Malware">

<meta name="DC.Subject" xml:lang="en" content="Risk">

<meta name="citation_keywords" xml:lang="en" content="Censorship">

<meta name="citation_keywords" xml:lang="en" content="Behavioural Advertising">

<meta name="citation_keywords" xml:lang="en" content="Malware">

<meta name="citation_keywords" xml:lang="en" content="Risk">

ويلاحظ أنه في تاج الميتا للكلمات المفتاحية تم الفصل بين هذه الكلمات بشبه شارحة، وهذا خلافا لما يوصى به باستخدام الفاصلة comma للفصل بين الكلمات المفتاحية. أما نسبة المقالات التي ورد بها عنصر الكلمات المفتاحية بشكل لا يتناسب مع المقالات، فقد قدرت بنسبة 33.33% (3 مقالات) من إجمالي المقالات التي ورد بها هذا العنصر. وقد انتمت هذه المقالات إلى دوريةD-Lib Magazine والتي أدرجت كلمات مفتاحية تتناسب مع الدورية نفسها وليس مع كل مقالة بها كالتالي:

<title>Access to Research Data</title>

<meta name="keywords" content="D-Lib Magazine, Digital Libraries, Digital Library Research">

وهذا ما قامت به دورية Cybrarians Journal في عنصر الكلمات المفتاحية الواردة بمقالاتها.

- يلاحظ ارتفاع جودة قيم باقي عناصر ما وراء البيانات وذلك في إجمالي المقالات التي ورد بها كل عنصر من هذه العناصر، وتشمل هذه العناصر كل من المؤلف، والتاريخ، والمصدر / عنوان الدورية، واالعدد، والترقيم الدولي الموحد للدورية، واللغة، والمحدد، والمعرف الرقمي للكيان، ومحدد ملخص المقالة، ونوع المصدر، والحقوق، والتغطية، والمساهم فكل عنصر من هذه العناصر وردت قيم صحيحة له في نسبة 100% من إجمالي المقالات الوارد بها هذا العنصر. وهذا يدل على حرص هذه الدوريات على إدراج عناصر ما وراء البيانات، وبجودة عالية والدليل على ذلك أن هذه العناصر تم تسجيلها وفقا لأكثر من خطة مثل خطة بؤرة دبلن، أو لعناصر ما وراء البيانات لجوجل الباحث العلمي، كما يفسر جودة هذه العناصر أن هذه الدوريات تستخدم برامج إدراة محتوى أو تصميم مواقع مثل دروبال Drupal المستخدم في دورية Ariadne، ونظام الدوريات المفتوح OJS الذي سبقت الإشارة إليه، وهذه البرامج توفر عدد من البرامج الإضافية plugins التي تتيح إمكانات إضافة عناصر ما وراء البيانات آليا، وما على مصمم الموقع سوى إدخال القيم الصحيحة الخاصة بهذه المقالات.

- يلاحظ أن عنصر المصدر source الوارد في مقالات دورية First Monday استخدم في إدراج بيانات الدورية مثل عنوان الدورية التي تنتمي إليها هذه المقالة، والترقيم الدولي الموحد لها، والموقع الخاص بهذه الدورية كالآتي:

<meta name="DC.Source" content="First Monday">

<meta name="DC.Source.ISSN" content="13960466">

<meta name="DC.Source.URI" content="http://firstmonday.org/ojs/index.php/fm">

وهذا هو نفس ما اتبعته مجلة المكتبات والمعلومات التي استخدمت عنصر المصدر source، ولذلك فعلى الرغم من أن القيم الواردة في هذا العنصر مناسبة لكل مقالة، إلا أن استخدام العنصر نفسه يعد غير صحيحا لإدراج بيانات الدورية، والأصح هو استخدام عنصرجزء من IsPartOf وهذا ما سبق الإشارة إليه.

من هذه النتائج وبشكل إجمالي وعلى الرغم من قلة عدد مقالات الدوريات الإلكترونية الأجنبية مقارنة بمقالات الدوريات الإلكترونية العربية التي تم دراستها، فإنه يمكن استنتاج أن جودة ما وراء البيانات في مقالات الدوريات الإلكترونية الأجنبية حققت مستوى متقدم عن مقالات الدوريات الإلكترونية العربية.

خلاصة الفصل:

تتمثل أبرز النتائج التي توصل لها هذا الفصل في الآتي:

1- توافرت عناصر ما وراء البيانات - وذلك بخلاف عنصر العنوان- في جميع مقالات الدوريات الإلكترونية العربية في مجال المكتبات والمعلومات بنسبة 100% من إجمالي المقالات، كما توافرت كذلك عناصر ما وراء البيانات في جميع المقالات للدوريات الإلكترونية الأجنبية في مجال المكتبات والمعلومات.

2- كانت أكثر خطط ما وراء البيانات استخداما في إدراج عناصر ما وراء البيانات في مقالات الدوريات الإلكترونية سواء العربية منها أو الأجنبية في المجال هي تيجان الميتا HTML Meta Tags، بينما استخدمت بؤرة دبلن في نسبة صغيرة (2.7%) من مقالات الدوريات الإلكترونية العربية انتمت إلى دورية واحدة فقط هي دورية مجلة المكتبات والمعلومات، وهذا يتقارب مع مقالات الدوريات الإلكترونية الأجنبية حيث جاءت هذه الخطة في المرتبة الأخيرة من بين خطط ما وراء البيانات المستخدمة في إدراج عناصر ما وراء البيانات وفي مقالات دورية واحدة فقط هي دورية First Monday.

3- انفردت مقالات الدوريات الإلكترونية الأجنبية بإدراج عناصر ما وراء البيانات وفقا لتيجان هاي ويربريس Highwire Press tags التي يوصي بها جوجل الباحث العلمي Google Scholar والتي استخدمت في نسبة 66.67% من المقالات محل الدراسة وقد انتمت هذه المقالات إلى دوريتين هما Ariadne، وFirst Monday.

4- بلغ أقصى عدد من عناصر ما وراء البيانات التي وردت في مقالات الدوريات الإلكترونية العربية في المجال (11) عنصرا، وردت في مقالات مجلة المكتبات والمعلومات، بينما بلغ أقصى عدد من عناصر ما وراء البيانات في مقالات الدوريات الإلكترونية الأجنبية في المجال (18) عنصرا وذلك في مقالات دورية First Monday.

5- بلغ إجمالي أنواع عناصر ما وراء البيانات المستخدمة في إجمالي مقالات الدوريات الإلكترونية العربية في المجال (13) عنصرا هي: العنوان title، والوصف description، والكلمات المفتاحية keywords /أو الموضوع subject، والمؤلف author/ أو المنشيء creator، والتاريخ date، واللغة language، والمحدد identifier، والمصدر source، ونوع المحتوى content-type، والمولد generator، والروبوت robot، والعرض viewport، ومدى التوافق X-UA-compatible.

6- بلغ إجمالي عدد عناصر ما وراء البيانات الواردة في مقالات الدوريات الإلكترونية الأجنبية (20) عنصرا تشمل العنوان، والوصف، والكلمات المفتاحية / الموضوع، والمؤلف، والتاريخ، والمصدر، والعدد issue، والمجلد volume، والترقيم الدولي الموحد للدورية ISSN، واللغة، والمحدد، والمعرف الرقمي للكيان DOI، وموقع ملخص المقالة citation_abstract_html_url، ونوع المصدر type، والحقوق rights، والتغطية coverage، والمساهم contributor، ونوع المحتوى/ الشكل، والمولد، ورقم إصدارة عناصر ما وراء البيانات لجوجل الباحث العلمي gs_meta_revision.

7- اشتركت كل من مقالات الدوريات العربية والأجنبية في مجال المكتبات والمعلومات في عشرة (10) عناصر شملت: العنوان، والوصف، والكلمات المفتاحية / الموضوع، والمؤلف / المنشئ، والتاريخ، واللغة، والمحدد، والمصدر، ونوع المحتوى، والمولد. بينما انفردت مقالات الدوريات الإلكترونية الأجنبية في فئة عناصر ما وراء البيانات الوصفية بتسعة عناصر لم ترد في مقالات الدوريات الإلكترونية العربية هي: العدد، والمجلد، والترقيم الدولي الموحد للدورية، والمعرف الرقمي للكيان، وموقع ملخص المقالة، ونوع المصدر، والحقوق، والتغطية، والمساهم.

8- كانت أكثر العناصر (التيجان) ورودا في مقالات الدوريات الإلكترونية العربية في المجال هي: العنوان، والوصف، والكلمات المفتاحية، ونوع المحتوى حيث ورد كل منها في جميع مقالات الدوريات الإلكترونية العربية محل الدراسة (بنسبة 100%)،

وباستثناء عنصر الكلمات المفتاحية فقد كانت هذه العناصر هي نفسها أكثر العناصر ورودا في مقالات الدوريات الإلكترونية الأجنبية بنسبة 100% من إجمالي مقالاتها: العنوان، والوصف، ونوع المحتوى.

9- لم تحقق عناصر ما وراء البيانات في الدوريات العربية في مجال المكتبات والمعلومات مستوى متقدم من الجودة، حيث كانت أغلب العناصر لا تتناسب مع كل مقالة من المقالات التي ترد بها، وذلك على عكس مقالات الدوريات الإلكترونية الأجنبية التي حققت مستوى متقدم من الجودة في عناصر ما وراء البيانات في مقالاتها.

10- كانت أكثر الدوريات الإلكترونية العربية في مجال المكتبات والمعلومات اهتماما بعناصر ما وراء البيانات وجودتها هي مجلة المكتبات والمعلومات سواء من حيث عدد العناصر المسجلة لكل مقالة، أو الخطط المستخدمة في ترميز هذه العناصر، أو حتى جودة القيم الواردة بها.

11- كانت أكثر الدوريات الإلكترونية الأجنبية في مجال المكتبات والمعلومات اهتماما بعناصر ما وراء البيانات وجودتها دورية First Monday سواء من حيث عدد عناصر ما وراء البيانات بها، أو من حيث خطط ما وراء البيانات المستخدمة، أو جودة قيم العناصر الواردة بها.

الفصل الثالث

قياس تأثير ما وراء البيانات على كفاءة استرجاع مقالات الدوريات الإلكترونية العربية في مجال المكتبات والمعلومات

ويشتمل هذا الفصل على النقاط التالية:

- اختبار تأثير ما وراء البيانات على كفاءة استرجاع مقالات الدوريات الإلكترونية العربية في مجال المكتبات والمعلومات من محركات البحث.
- إمكانية استثمار ما وراء البيانات في مقالات الدوريات الإلكترونية العربية في مجال المكتبات والمعلومات.

الفصل الثالث

قياس تأثير ما وراء البيانات على كفاءة استرجاع مقالات الدوريات الإلكترونية العربية في مجال المكتبات والمعلومات

تمهيد:

يسعى هذا الفصل إلى اختبار تأثير ما وراء البيانات وتحديدا ما وراء البيانات المكودة باستخدام عناصر بؤرة دبلن Dublin Core، أو مفردات سكيما Schema.org باستخدام إطار وصف المصدر في الخصائص RDFa على كفاءة استرجاع صفحات العنكبوتية بصفة عامة، ومقالات الدوريات الإلكترونية العربية في مجال المكتبات والمعلومات بصفة خاصة من ثلاثة محركات بحث هي: جوجل Google، وياهو Yahoo، وبينج Bing. واستخدم في ذلك المنهج التجريبي حيث تم إعداد تجربة على عينة من مقالات الدوريات الإلكترونية العربية في مجال المكتبات والمعلومات.

كما يحاول هذا الفصل التحقق من إمكانية استثمار ما وراء البيانات الواردة في مقالات الدوريات الإلكترونية العربية في مجال المكتبات والمعلومات في إعداد تسجيلات ببليوجرافية

وفقـا لشـكل الاتصـال مـارك 21 MARC21 يمكـن إيـداعها في فهـارس المكتبـات أو أحـد المستودعات الرقمية، أو في إعداد كشاف لهذه المقالات.

1 اختبار تأثير ما وراء البيانات على كفاءة استرجاع مقالات الدوريات الإلكترونيـة العربيـة في مجال المكتبات والمعلومات من محركات البحث:

لاختبار تأثير ما وراء البيانات على كفاءة استرجاع مقالات الدوريات الإلكترونية العربيـة في المجال تم إعداد تجربة لتحديد مدى صحة الفرض:

- لا تؤثر ما وراء البيانات عـلى قـدرة محركـات البحـث عـلى اسـترجاع مقـالات الـدوريات الإلكترونية العربية في مجال المكتبات والمعلومات.
- وقد تم قياس هذا الفرض بالنسبة لكل محرك بحث على حده.

1-1 مراحل التجربة:

1-1-1 مراسلة الدوريات عينة الدراسة:

تمـت مراسـلة كـل دوريـة مـن الـدوريات الخمسـة عينـة الدراسـة عـن طريـق البريـد الإلكتروني (Cybrarians Journal، والعربيـة 3000، ودراسـات المعلومـات، والمجلـة العربيـة للدراسات المعلوماتية، ومجلة المكتبات والمعلومات) من أجل أخـذ إذن إعـادة نشر المقـالات الصادرة عنها على الإنترنت عبر مواقع ستخصص للتجربة، وتم ذلك حفاظا على حقوق الملكية الفكرية لهذه الدوريات، وتحسبا لوقف التجربة نتيجة انتهاك هذه الحقوق. وقـد اسـتجابت جميع الدوريات فيما عدا دورية واحدة هي مجلـة المكتبـات والمعلومـات، ولـذلك فقـد تـم استبعاد مقالاتها من العينة التي خضعت للتجربة.

2-1-1 تحديد عينة التجربة من المقالات:

تم اختيار عينة التجربة من المقالات عشوائيا، وذلك بمعدل مقالـة واحـدة ممثلـة لكـل عدد من الأعداد المتاحة على موقع كل دورية حتى يناير 2015، على أن تكون أول مقالة تـرد بعد الافتتاحية إن وجدت ليصبح إجمالي عينة المقالات التي خضعت للتجربة (72) مقالة.

ويوضح الجدول التالي رقم (3-1) عينة التجربة من المقالات:

الجدول رقم (3-1) عينة التجربة من مقالات الدوريات الإلكترونية العربية في مجال المكتبات والمعلومات

الدورية	الأعداد المتاحة حتى يناير 2015	عينة التجربة من المقالات	نصيب كل دورية من المقالات في عينة التجربة
Cybrarians Journal	36	36	%50
العربية 3000	17	17	%23.61
دراسات المعلومات	14	14	%19.44
المجلة العربية للدراسات المعلوماتية	5	5	%6.95
الإجمالي	72	72	%100

يتضح من الجدول السابق رقم (3-1) ما يلي:

- بلغ إجمالي عدد المقالات التي خضعت للتجربة (72) مقالة من إجمالي المقالات الصادرة عن هذه الدوريات والبالغ عددها (647) مقالة، ووفقا لبرنامج Sample Size Calculator [1] تبين أن درجة الثقة في هذه العينة تبلغ 90%، وبهامش خطأ 10.90 %.
- كان النصيب الأكبر من عينة المقالات التي خضعت للتجربة منشورة في دورية cybrarians Journal حيث كان عدد المقالات التي خضعت للتجربة منها (36) مقالة بنسبة حوالي 50% من إجمالي عينة التجربة، تلتها العربية 3000 بـ(17) مقالة بنسبة تقدر بنحو 23.61%، لترد بعد ذلك دراسات المعلومات بعدد (14) مقالة بنسبة بلغت 19.44%، وأخيرا جاءت المجلة العربية للدراسات المعلوماتية بعدد (5) مقالات وبنسبة 6.95%. ويرجع اختلاف مساهمة كل دورية من حيث عدد المقالات التي خضعت للتجربة لاختلاف عدد الأعداد الصادرة عن كل منها.

---

(1) Raosoft. Sample Size Calculator by Raosoft, Inc.[online]. Available at: http://www.raosoft.com/samplesize.html [Accessed 24/2/2015].

3-1-1 تجميع عينة المقالات:

تم تجميع عينة المقالات وحفظها على الحاسب الآلي.

4-1-1 تصميم التجربة:

تم تصميم التجربة وفقا لتصميم التجربة البعدية، لتشمل ثلاث مجموعات تحتوي كل منها على نفس العينة من المقالات (72 مقالة) وفقا لما يلي:

أ - المجموعة الضابطة - المجموعة (أ):

هي المجموعة التي لا تحتوي مقالاتها على ما وراء البيانات. حيث تم إعداد صفحات جديدة بلغة ترميز النص الفائق HTML لهذه العينة من المقالات وفقا للإصدارة الخامسة HTML5، وتمت الاستعانة في إعداد هذه الصفحات بأحد البرامج مفتوحة المصدر لتصميم وتحرير المواقع وهو برنامج بلوجريفون BlueGriffon[1] الإصدارة 1.7.2.

ب- المجموعة التجريبية الأولى - المجموعة (ب):

هي المجموعة التجريبية التي أدخل عليها متغير التجربة وهو ما وراء البيانات وفقا لبؤرة دبلن Dublin Core. واتبع في ذلك الخطوات التالية:

ب1- أخذ نسخة من صفحات المقالات السابق إعدادها بلغة HTML الإصدارة الخامسة في المجموعة الضابطة وإضافة عناصر ما وراء البيانات في شفرة المصدر source code لكل مقالة، وفقا لبؤرة دبلن تبعا لما يلي:

ب1-1 استخدام مفردات بؤرة دبلن لما وراء البيانات DCTERMS حيث أنها الأحدث كما أنها تشتمل على جميع العناصر الخاصة ببؤرة دبلن سواء البسيطة منها أو ما كان يعرف منها سابقا بالمقيدة أو المحددة. وهذه المفردات يتم التعبير عنها بالبادئة DCTERMS. وتم الإعلان عنها في شفرة مصدر Source Code الصفحات التي تم إعدادها من خلال الترميز التالي:

```
<link rel="schema.DCTERMS" href="http://purl.org/dc/terms/">
```

(1) BlueGriffon [online]. Available at: http://bluegriffon.org/ .[Accessed 8/11/2014].

ب1-2 تحديد عناصر بؤرة دبلن لما وراء البيانات التي يمكن استخدامها في وصف المقالات، ونظرا لأن جميع عناصر بؤرة دبلن عناصر اختيارية، فقد تم اختيار عشرة عناصر فقط من مفردات بؤرة دبلن DC Terms لوصف كل مقالة، وهي:

العنوان Title، المنشئ Creator، الموضوع Subject، الوصف Description، الناشر Publisher، المساهم (إن وجد) Contributor، التاريخ Date، المحدد Identifier، جزء من IsPartOf، الحقوق Rights. وتم الاكتفاء بهذه العناصر لوصف مقالات الدوريات ذلك لأنها تستطيع توفير وصفا كافيا لمقالات الدوريات. وهو ما ينفي أحد الفروض التي وضعتها الدراسة والذي ينص على أنه: "يعجز بؤرة دبلن عن وصف مقالات الدوريات الإلكترونية العربية في المجال بشكل مناسب".

وفيما يلي تفصيل لكيفية استخدام كل عنصر وطريقة ترميزه.

1 العنوان Title	
ممارسة التطبيق	تسجيل عنوان المقالة، والعنوان الفرعي إن وجد.
طريقة الترميز	استخدم عنصر الميتا meta مع كل من الخاصية name التي سجل فيها اسم عنصر ما وراء البيانات مسبوقا بالبادئة DCTERMS التي تعبر عن استخدام مفردات بؤرة دبلن، وسجلت القيمة أو عنوان المقالة نفسه في الخاصيةcontent كما يلي: <meta name="DCTERMS.Title" content="مكتبة المدينة الرقمية: الواقع والمستقبل">
2 المنشئ Creator	
ممارسة التطبيق	تسجيل اسم المؤلف كاملا بشكله الطبيعي، وفي حالة وجود أكثر من مؤلف تم تكرار العنصر.
طريقة الترميز	مثال: <meta name="DCTERMS.Creator" content="عزة فاروق جوهري"> <meta name="DCTERMS.Creator" content="أريج الحازمي">

3 الموضوع Subject	
ممارسة التطبيق	تسجيل الكلمات أو الجمل المفتاحية الحرة للتعبير عن موضوع المقالات، ولم يتم استخدام مفردات قائمة رؤوس موضوعات مقيدة لأن قائمة رؤوس الموضوعات المقيدة لا تتناسب مع البيئة المفتوحة في الإنترنت التي لا تلتزم بقوائم رؤوس موضوعات ويمكن البحث فيها عن المصادر بكلمات مفتاحية حرة وهذا ما أشارت إليه تيلور[1] من أن اللغات المقيدة ليست بالضرورة ملائمة لجميع بيئات المعلومات؛ حيث أنها توظف بشكل أفضل في سياق النظم الشمولية أو الكبيرة contained system، وأن من السهولة إنشاءها وصيانتها وتطبيقها فقط في نظم كالكشافات، وقواعد البيانات، والفهارس والتي تعتمد على الخبرة البشرية، أما الطريقة المقننة للمصطلحات فليس بإمكانها وليس باستطاعتها أن تعمل بشكل جيد مع نظم كبيرة وموزعة مثل الإنترنت. ويتفق ذلك مع ما أشار إليه لانكستر[2] من أن اللغة الطبيعية غير المقيدة أكثر انتشارا وأكثر جدوى من غيرها مع استخدام الحاسبات الإلكترونية في استرجاع المعلومات. وقد تم استقاء الكلمات المفتاحية من العنوان، والنص الكامل للمقالة. وتجدر الإشارة إلى أنه تم استخدام اللغة العربية في التعبير عن موضوع المقالات المكتوبة باللغة العربية، أما في حالة المقالات المكتوبة باللغة الإنجليزية فقد تم استخدام اللغة الإنجليزية في التعبير عن موضوعات هذه المقالات بالإضافة إلى تسجيل كلمات مفتاحية باللغة العربية. وفي حالة وجود أكثر من كلمة مفتاحية، تم تكرار العنصر.
طريقة الترميز	<meta name="DCTERMS.Subject" content="مكتبة المدينة الرقمية"> <meta name="DCTERMS.Subject" content="المكتبات الرقمية"> <meta name="DCTERMS.Subject" content="السعودية">

(1) تيلور، ألين ج. تنظيم المعلومات. جزء 1. / تأليف ألين ج. تيلور، دانيل ن. جودري ؛ ترجمة هاشم فرحات. الرياض: جامعة الملك سعود، النشر العملي والمطابع،2012. ص 510.

(2) لانكستر، ولفرد. نظم استرجاع المعلومات/ ترجمة حشمت قاسم. القاهرة: دار غريب، 1981. ص 382.

4 الوصف Description	
ممارسة التطبيق	تسجيل ملخص عن المقالة، وفي أغلب الأحوال قامت الباحثة بوضع هذا الملخص، ليكون مختصرا ومعبرا قدر الإمكان.
طريقة الترميز	<meta name="DCTERMS.Description" content="تتناول الدراسة المكتبة الرقمية مع التركيز على مكتبة المدينة الرقمية كنموذج لهذا النوع من المكتبات، فتستعرض ماهية المكتبة الرقمية وأهميتها وبنيتها وتنظيمها ومواردها وتجهيزاتها وخدماتها لاستشراف المستقبل لهذا النمط من المكتبات" >
5 الناشر Publisher	
ممارسة التطبيق	تسجيل ناشر الدورية في هذا العنصر مع كل مقالة.
طريقة الترميز	<meta name="DCTERMS.Publisher" content="البوابة العربية للمكتبات والمعلومات">
6 المساهم (إن وجد) Contributor	
ممارسة التطبيق	تم استخدام هذا العنصر لتسجيل المساهم في المقالة، مثل المترجم.
طريقة الترميز	<meta name="DCTERMS.Contributor" content="حسن محمد خليل"> <meta name="DCTERMS.Contributor" content="عزة محمود علي حسن">
7 التاريخ Date	
ممارسة التطبيق	سُجل تاريخ نشر العدد الذي ظهرت فيه المقالة، وفقا لصيغة ISO 8601 التي يُوصى باستخدامها. وفي حالة الدوريات التي تصدر مقالاتها وفقا للتاريخ الهجري والميلادي، تم الاكتفاء بتسجيل التاريخ الميلادي فقط.
طريقة الترميز	<meta name="DCTERMS.Date" content="2008-03">

8 المحدد Identifier	
ممارسة التطبيق	تسجيل الموقع الموحد للمصدر URL الخاص بالمقالة.
طريقة الترميز	نظرا لأن قيمة العنصر عبارة عن موقع موحد للمصدر URL، وليست قيمة حرفية، كما في العناصر السابقة، فقد تم استخدام العنصر link مع كل من الخاصية rel، والخاصية href كالتالي: <link rel="DCTERMS.Identifier" href="http://ajainlis.com/cybrarians-journal-articles/-مكتبة-المدينة-الرقمية-الواقع-والمستقبل.html">
9 جزء من IsPartOf	
ممارسة التطبيق	تنص إرشادات بؤرة دبلن حول كيفية تسجيل هذا العنصر، بأن تسجل بيانات هذا العنصر إما بأن تكون الموقع URL للكيان الذي يشكل المصدر الموصوف جزءا منه وهو هنا موقع العدد الخاص بالدورية، أو أن يسجل نص بسيط يعبر عن هذا الكيان. وقد اختارت الباحثة تكرار هذا العنصر بحيث يسجل مرة في شكل موقع URL، ويسجل مرة أخرى كقيمة نصية لتسجيل بيانات العدد وعنوان الدورية التي وردت بها المقالة.
طريقة الترميز	<meta name="DCTERMS.IsPartOf" content="cybrarians journal العدد 15"> <link rel="DCTERMS.IsPartOf" href="http://www.journal.cybrarians.org/index.php?option=com_sectionex&view=category&id=15&Itemid=57">
10 الحقوق Rights	
ممارسة التطبيق	تسجيل الجهة التي تمتلك حقوق النشر لمقالات الدورية.
طريقة الترميز	<"meta name="DCTERMS.Rights" content="cybrarians journal>

ب1-3 التأكد من صحة الترميز لمقالات هذه المجموعة من خلال الاستعانة بمصحح لغة ترميز النص الفائق HTML المتاح من خلال اتحاد العنكبوتية W3C [1]، و تـم إدراج ترميـز جزء الرأس <head> الموجود به عناصر ما وراء البيانات وفقـا لبـؤرة دبلـن في هـذا المصـحح، وأشار المصحح دائما إلى وجود خطأ واحد في عنصر المحدد Identifier كما في الشكل رقم (3-1) حيث تمت الإشارة إلى عدم صحة استخدام هذا العنصر مع الخاصية rel.

1. Error | Bad value DCTERMS.Identifier for attribute rel on element link: The string dcterms.identifier is not a registered keyword.

From line 18, column 5; to line 18, column 134

```
-03">↩    <link rel="DCTERMS.Identifier" href="http://ajainlis.com/cybrarians-journal-articles/مكتبة-المدينة-الرقمية-الواقع-والمستقبل.html">↩    <
```

Syntax of list of link-type keywords:

A whitespace-separated list of link types listed as allowed on <link> in the HTML specification or listed as an allowed on <link> on the Microformats wiki without duplicate keywords in the list. **You can register link types on the Microformats wiki yourself.**

Document checking completed.

Source

```
1. <!DOCTYPE html>↩
2. <html dir="rtl" lang="ar">↩
3.   <head>↩
4.     <meta content="text/html; charset=utf-8" http-equiv="content-type">↩
5.     <title>مكتبة المدينة الرقمية : الواقع والمستقبل</title>↩
6.     <link rel="stylesheet" href="styles.css" type="text/css" media="screen">↩
7.     <link rel="schema.DCTERMS" href="http://purl.org/dc/terms/">↩
8.     <meta name="DCTERMS.Title" content="مكتبة المدينة الرقمية : الواقع والمستقبل">↩
9.     <meta name="DCTERMS.Creator" content="عزة فاروق جوهري">↩
10.    <meta name="DCTERMS.Creator" content="أريج الحازمي">↩
11.    <meta name="DCTERMS.Subject" content="مكتبة المدينة الرقمية">↩
12.    <meta name="DCTERMS.Subject" content="المكتبات الرقمية">↩
13.    <meta name="DCTERMS.Subject" content="المكتبة الرقمية">↩
14.    <meta name="DCTERMS.Subject" content="السعودية">↩
15.    <meta name="DCTERMS.Description" content="تتناول الدراسة المكتبة الرقمية مع التركيز على مكتبة المدينة الرقمية كنموذج لهذا النوع من المكتبات، فتستعرض ماهية المكتبة الرقمية وأهميتها وبنيتها وتنظيمها ومواردها وتجهيزاتها وخدماتها لاستشراف المستقبل لهذا النمط من المكتبات">↩
16.    <meta name="DCTERMS.Publisher" content="البوابة العربية للمكتبات والمعلومات">↩
17.    <meta name="DCTERMS.Date" content="2008-03">↩
18.    <link rel="DCTERMS.Identifier" href="http://ajainlis.com/cybrarians-journal-articles/مكتبة-المدينة-الرقمية-الواقع-والمستقبل.html">↩
19.    <meta name="DCTERMS.IsPartOf" content="cybrarians journal 15 العدد">↩
20.    <link rel="DCTERMS.IsPartOf" href="http://www.journal.cybrarians.org/index.php?option=com_sectionex&amp;view=category&amp;id=15&amp;Itemid=57">↩
21.    <meta name="DCTERMS.Rights" content="cybrarians journal">↩
```

الشكل رقم (3-1) مصحح لغة ترميز النص الفائق لاتحاد العنكبوتية W3C

إلا أن هذا يتعارض مع إرشادات بؤرة دبلن حول كيفية ترميـز هـذا العـنصر باسـتخدام لغـة ترميز النص الفائق والتي أوصت بأنـه عنـدما تكـون قيمـة البيـان عبـارة عـن محـدد موحـد

(1) The W3C Markup Validation Service [online]. Available at: https://validator.w3.org/

للمصدر URI ففي هذه الحالة يتم استخدام الخاصية rel بدلا من الخاصية name[1]. وهـذا من الممكن أن يكون مع أي عنصر من عناصر بؤرة دبلن كما أن الإصدارة الخامسـة مـن لغـة ترميز النص الفائق HTML 5 أتاحت القيام بذلك [2]. يضـاف إلى ذلـك أن محركـات البحـث يمكنها قراءة صفحات العنكبوتيـة بلغـة ترميـز الـنص الفـائق حتـى وإن كـان بهـا أخطـاء في ترميزها.

ج- المجموعة التجريبية الثانية - المجموعة (ج):

هي المجموعة التجريبية التي تشتمل على عينة المقالات مضاف إليها مـا وراء البيانـات وفقا لمفردات سكيما Schema.org وباسـتخدام إطـار وصـف المصـدر في الخصـائص RDFa. واتبع في ذلك الخطوات التالية:

ج1- أخذ نسخة من صفحات المقالات السابق إعدادها بلغة HTML الإصدارة الخامسة في المجموعة الضابطة وإضافة عناصر ما وراء البيانـات في شـفرة المصـدر source code لكـل مقالة، وفقا لمفردات سكيما Schema.org وهي المفردات التي تشكل جهـد وتعـاون مشـترك بين محركات البحث الرئيسية جوجل، وبينج (مايكروسوفت)، وياهو، ومحرك البحث الـروسي ياندكس Yandex وفقا لما يلي:

ج1-1 تحديد مفردات سكيما Schema.org التي استخدمت في وصف المقالات، وقد تم اختيار عشرة مفردات فقط تعبر عن عشرة عناصر لوصف كل مقالة، وهي:

العنــوان الــرئيسي headline، المؤلــف author، الكلــمات المفتاحيــة keywords، الوصـف description، النـاشر publisher، المسـاهم (إن وجـد) contributor، تـاريخ النشر datePublished، الموقع الموحد للمصدر url، جزء من isPartOf، مالـك الحقـوق

(1) Expressing Dublin Core metadata using HTML/XHTML meta and link elements [online]. Available at: http://dublincore.org/documents/dc-html/. [Accessed 18/8/2015].

(2) Existing rel values · Microformats Wiki. [online] Available at: http://microformats.org/wiki/existing-rel-values#HTML5_link_type_extensions [Accessed 24/6/2016].

copyright Holder. وتم الاكتفاء بهذه العناصر لوصف مقالات الدوريات لأنها تستطيع توفير وصفا كافيا لمقالات الدوريات. وهو ما ينفي أحد الفروض التي وضعتها الدراسة والذي ينص على أنه: " تعجز مفردات سكيما Schema.org عن وصف مقالات الدوريات الإلكترونية العربية في المجال بشكل مناسب".

ج1-2 تم ترميز هذه المفردات باستخدام إطار وصف المصدر في الخصائص RDFa، وهو عبارة عن مجموعة من الخصائص الإضافية التي يتم استخدامها في إدراج ما وراء البيانات وفقا لإطار وصف المصدر RDF في صفحات العنكبوتية المكتوبة بلغة ترميز النص الفائق HTML – انظر الفصل الأول – ومن بين هذه الخصائص تم استخدام الخصائص التالية:

Vocab, typeof, property, resource, content.

ج1-3 تم الإعلان عن استخدام مفردات سكيما Schema.org باستخدام الخاصية vocab التي يحدد فيها الموقع الموحد URL الخاص بهذه المفردات مع ضرورة تحديد نوع المصدر الموصوف في الخاصية typeof وهو هنا المقالات كما يلي:

```
<body>
…
  <article vocab="http://schema.org/" typeof="ScholarlyArticle">
…
  </article>
</body>
```

وتجدر الإشارة إلى نقطة هامة وهي أنه في حالة إدراج ما وراء البيانات باستخدام إطار وصف المصدر في الخصائص RDFa فإنه يمكن أن تدرج هذه البيانات في جزء الرأس head في عنصر الميتا meta باستخدام كلا من الخاصية property، والخاصية content كما يلي:

```
<head>
<meta property="author"  content="عزة فاروق جوهري">
…
</head>
```

وكذلك يمكن أن تدرج في جزء الجسم body لصفحة العنكبوتية وفي هذه الحالة يمكن أن تكون قيم عناصر ما وراء البيانات جزءا من البيانات المعروضة في الصفحة التي تظهر للمستفيد كما في المثال التالي:

```
<body>
…
<span property="author">عزة فاروق جوهري</span>
…
</body>
```

كما يمكن إخفاء قيمة عنصر ما وراء البيانات باستخدام الخاصية content.

وقد اتبعت الدراسة إدراج ما وراء البيانات في جزء الجسم body، نظرا لأن جوجل Google أشار إلى عدم دعمه لإدراج ما وراء البيانات في عنصر الميتا meta باستخدام كلا من الخاصية property، والخاصية content.[1]

وفيما يلي تفصيل لكيفية استخدام كل مفردة من مفردات سكيما Schema.org وطريقة ترميزها.

1 العنوان الرئيسي headline	
ممارسة التطبيق	تسجيل عنوان المقالة والعنوان الفرعي إن وجد
طريقة الترميز	<h1 property="headline">مكتبة المدينة الرقمية: الواقع والمستقبل</h1> يلاحظ هنا أن قيمة العنوان نفسه أي عنوان المقالة هو جزء من النص المعروض في الصفحة.
2 المؤلف author	
ممارسة التطبيق	تسجيل اسم المؤلف كاملا بشكله الطبيعي
طريقة الترميز	عزة فاروق جوهري

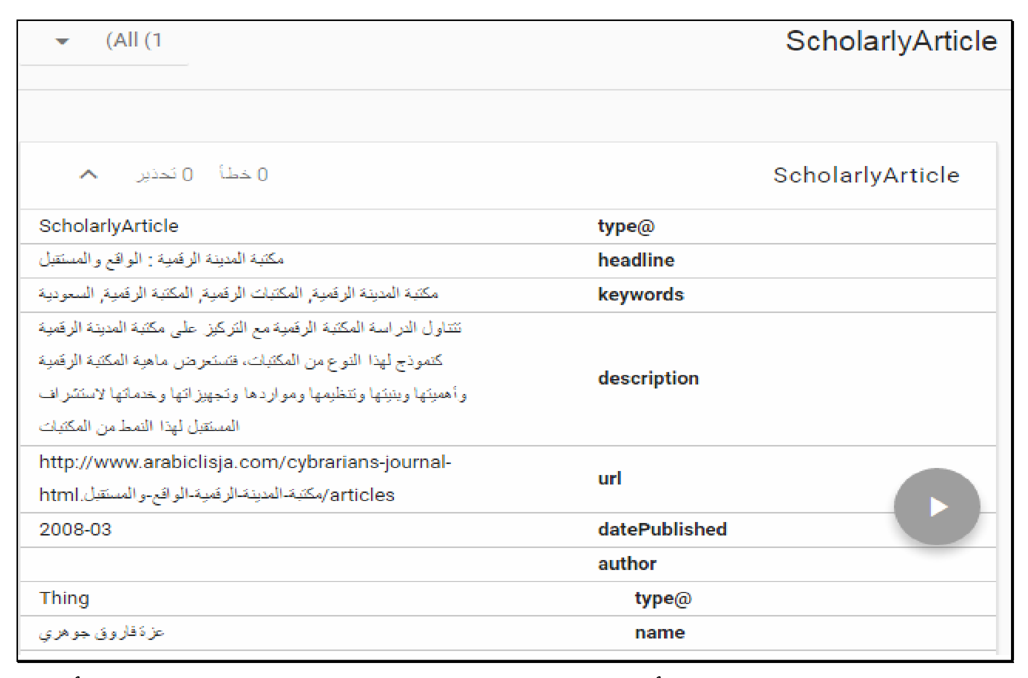
(1) Google Developers. Providing Structured Data. [online] Available at: https://developers.google.com/custom-search/docs/structured_data#using-meta-tags.[Accessed 22/10/2015].

3 الكلمات المفتاحية keywords	
ممارسة التطبيق	تسجيل الكلمات أو الجمل المفتاحية المعبرة عن محتوى المقالات وتم استقاءها من العنوان ونص المقالة، وتم الفصل بين الكلمات المفتاحية بفاصلة comma
طريقة الترميز	تم استخدام الخاصيةcontent لتسجيل الكلمات المفتاحية وهذا يعني إخفاء هذه القيمة وعدم ظهورها للمستفيد، نظرا لأن هذه الكلمات المفتاحية لم تكن جزءا من نص المقالة الأصلي.
4 الوصف description	
ممارسة التطبيق	سُجل في هذا العنصر ملخص عن المقالة معبرا ومختصرا قدرا الإمكان.
طريقة الترميز	
5 الناشر publisher	
ممارسة التطبيق	تسجيل ناشر الدورية.
طريقة الترميز	
6 المساهم (إن وجد) contributor	
ممارسة التطبيق	تسجيل المساهم مثل المترجم أو أي دور آخر بخلاف المؤلف.
طريقة الترميز	حسن محمد خليل
7 تاريخ النشر datePublished	
ممارسة التطبيق	تاريخ نشر العدد الذي ظهرت فيه المقالة وفقا لصيغة ISO 8601 المُوصى باستخدامها.

طريقة الترميز	 يونيـو 2005
8 الموقع الموحد للمصدر url	
ممارسة التطبيق	تسجيل الموقع الموحد للمصدر URL المتاحة عليه المقالة
طريقة الترميز	تم استخدام الخاصية resource لتسجيل الموقع الموحد للمقالة كالتالي:
9 جزء من isPartOf	
ممارسة التطبيق	تسجيل بيانات العدد والدورية التي تنتمي إليها المقالة.
طريقة الترميز	 cybrarians journal. العدد 5
10 مالك الحقوق copyrightHolder	
ممارسة التطبيق	تسجيل الجهة المالكة لحقوق نشر مقالات الدورية
طريقة الترميز	

ج1-4 تم التأكد من صحة الترميز لمقالات هذه المجموعة باستخدام أداة جوجل Structured Data Testing Tool [1] والتي أشارت إلى عدم وجود أخطاء كما في الشكل رقم (2-3):

(1) أداة اختبار البيانات المنظمة. متاح على:
https://search.google.com/structured-data/testing-tool



الشكل رقم (3-2) نتيجة اختبار أداة جوجل Structured Data Testing Tool لترميز أحد صفحات المجموعة (ج)

كما تم اختبار صحة هذا الترميز على أداة RDFa play[1] وتبين عدم وجود أخطاء كما يتضح من الشكل التالي رقم (3-3).

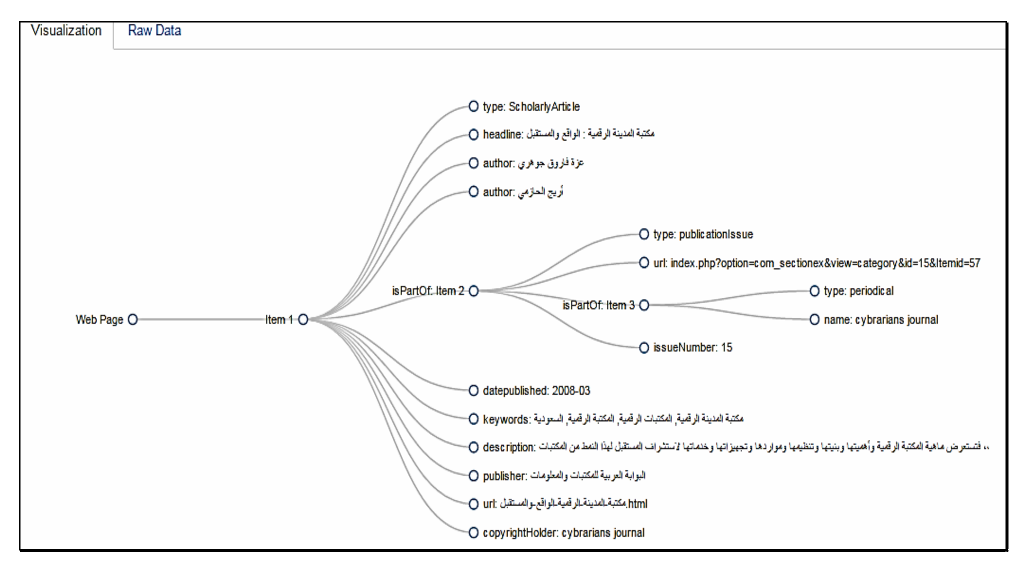
5-1-1 إجراء التجربة:

تم إجراء التجربة وفقا للخطوات التالية:

1- إنشاء ثلاثة مواقع في نفس الوقت (16 يناير 2016) لتتاح عليها مجموعات التجربة على الإنترنت وكانت كالتالي:

- www.lisaja.com: موقع لإتاحة مقالات المجموعة الضابطة - المجموعة (أ).

(1) RDFa play [online]. Available at: https://rdfa.info/play/



الشكل رقم (3-3) نتيجة اختبار ترميز أحد صفحات المجموعة (ج) على أداة RDFa play

- www.ajainlis.com: موقع لإتاحة مقالات المجموعة التجريبية الأولى - المجموعة (ب).
- www.arabiclisja.com: موقع لإتاحة مقالات المجموعة التجريبية الثانية - المجموعة (ج).

2- إيداع Submit المواقع في محركات البحث:

لضمان تكشيف المواقع المتاحة عليها مجموعات الدراسة تم إيداع أو إضافة Submit هذه المواقع في محركات البحث عينة الدراسة في نفس الوقت، وتطلب ذلك عمل حساب Account لكل موقع في أدوات مشرفي المواقع في كل من محرك البحث جوجل Google، ومحرك البحث بينج Bing. وتجدر الإشارة إلى أنه لم يتم عمل حساب خاص بمحرك البحث ياهو Yahoo نظرا لأن ذلك يتم من خلال أدوات مشرفي المواقع لبينج.

كما تمت إضافة ملف خريطة كل موقع sitemap.xml لهذه المحركات كخطوة إضافية لضمان تكشيف هذه المواقع.

3- إتاحة فترة زمنية أمام محركات البحث لتكشيف مواقع التجربة: وقد بلغت هذه الفترة شهرين، وتم مراجعة حساب أدوات مشرفي المواقع في كل محرك بحث حيث يتم فيها عرض حالة تكشيف الموقع. وتجدر الإشارة إلى تفوق محرك البحث جوجل في سرعة تكشيف هذه المواقع.

4- البحث عن كل مقالة من المقالات عينة التجربة في كل من محرك البحث جوجل، وياهو، وبينج وفقا لعنصرين أساسيين هما العنوان، والكلمات المفتاحية. حيث يعكس البحث بالعنوان قدرة محرك البحث على الاسترجاع اعتمادا على تكشيف النص الكامل الوارد في الموقع أو الصفحة، أما الكلمات المفتاحية فتعكس قدرة محرك البحث على الاسترجاع وفقا لتكشيف النص الكامل وكذلك الدلالة أو المعنى للموضوع أو الكلمات التي يتم البحث عنها خاصة إذا لم ترد في النص. وتم مراعاة النقاط التالية في عملية البحث:

عند البحث بالعنوان تم استخدام علامات التنصيص (" ") حتى تقتصر النتائج على نفس الكلمات وبنفس الترتيب، بما يؤدي إلى تقليل حجم النتائج المسترجعة، مع ضمان ظهور صفحات المجموعات التجريبية ضمن المئات الأولى من نتائج البحث المسترجعة.

للبحث بالكلمات المفتاحية تم إعداد استراتيجية تتناسب مع كل مقالة، وعندما يتطلب الأمر البحث بأكثر من كلمة مفتاحية تم الربط بينهما بعلامة الزائد مثال: برامج التعليم الببليوجرافي +المكتبات الأكاديمية.

5- تجميع النتائج من محركات البحث وطباعتها:

تم البحث عن المقالات وفي حالة العثور على المقالة، تم تحديد المجموعة التي تنتمي إليها وفقا لموقعها، مع تحديد ترتيب ظهورها في قائمة نتائج محرك البحث، ما بين ظهرت أولا، أو ثانيا، أو ثالثا. وعلى مقياس من 1 إلى 3 تم إعطاء المجموعة التي تظهر أولا في قائمة نتائج محركات البحث أعلى درجة على هذا المقياس أي (3)، والتي تظهر ثانيا (2)، والتي تظهر أخيرا أخذت أقل درجة على هذا المقياس أي (1). وتجدر الإشارة

إلى أنه تم التوقف عن البحث عند الصفحة رقم 15 من صفحات نتائج محركات البحث – أي تم التوقف عند النتيجة رقم 150 – وهذا مدى كاف فمعظم المستفيدين عادة ما تفحص أو تراجع أول عشرة مواقع في قائمة نتائج محرك البحث وأن 1 % فقط منهم يتحقق أو يراجع ما بعد الصفحة الثالثة من قائمة نتائج محرك البحث[1].

ونظرا لأن هذه المجموعات التي تم إنشاءها تعد محتوى متكرر، حيث أن كل مجموعة تحتوي على نفس العينة من المقالات وهذه المقالات في حد ذاتها نسخة مكررة من مقالات الدوريات الأصلية المتاحة على الإنترنت، فإن هذا يشكل أحد قيود هذه التجربة، حيث يحاول كل محرك من محركات البحث تقديم نتائج لا تمثل محتوى متكرر أو منسوخ وتقديم مواقع ذات محتوى فريد، وينتهج كل محرك بحث طريقة في تفادي ذلك، فمحرك البحث جوجل على سبيل المثال يقوم بعدم إظهار هذه التكرارات ويمكن الاطلاع عليها ضمن قائمة نتائج البحث عند الضغط على رابط "تكرار البحث باستخدام تضمين نتائج البحث التي تم حذفها" لإظهار النتائج المكررة كما يتضح من الشكل رقم (3-4)، ولذلك ففي محرك البحث جوجل كان يتم الضغط على هذا الرابط حتى تظهر جميع مواقع التجربة، أما في ياهو وبينج فكلا منهما لا يتيح هذه الإمكانية.

لعرض النتائج الأكثر صلة بالموضوع، حذفنا بعض الإدخالات الشبيهة 10 بالنتائج المعروضة فعلاً.
إذا كنت ترغب في ذلك، يمكنك تكرار البحث باستخدام تضمين نتائج البحث التي تم حذفها.

الشكل رقم (3-4) رسالة محرك البحث جوجل لإظهار النتائج المكررة التي تم حذفها من قائمة نتائج البحث

(1) Zhang, Jin and Dimitroff, Alexandra. The Impact of Webpage Content Characteristics on Webpage Visibility in Search Engine Results [online]. Part I. Information Processing and Management. Vol. 41 (2005). p 665. Available at: https://www.semanticscholar.org/paper/The-impact-of-webpage-content-characteristics-on-Zhang-Dimitroff/eb943e7a6307a85d851c3cc23026dcab1151ccac/pdf. [Accessed 28/9/2015].

2-1 نتائج التجربة:

تم استخدام رتبة الظهور لمقالات كل مجموعة في قائمة نتائج محركات البحث عينة الدراسة كمقياس لتحديد كفاءة الاسترجاع من هذه المحركات، فالمجموعة التي تظهر مقالاتها أولا في قائمة نتائج محرك البحث تحقق أداء أفضل من المجموعة التي تظهر مقالاتها في المرتبة الثانية أو الثالثة. و بعد تجميع البيانات تم إدخالها في برنامج الحزمة الإحصائية للعلوم الاجتماعية Statistical Package for Social Sciences (SPSS) للحصول على التحليل الإحصائي الوصفي كما تم الاستعانة بهذا البرنامج في إجراء أحد الاختبارات الإحصائية وهو اختبار كروسكال ويلز Kruskal Wallis Test الذي يختبر أكثر من عينتين مستقلتين بما يتناسب مع اشتمال هذه التجربة على ثلاث مجموعات، وتم حساب معنوية الاختبار على أساس إذا كان مستوى المعنوية أقل من (0.05) دل ذلك على وجود فروق ذات دلالة إحصائية بين مجموعات التجربة ورفض الفرض الصفري، وإذا كان مستوى المعنوية أكبر من (0.05) دل ذلك على عدم وجود فروق ذات دلالة إحصائية وقبول الفرض الصفري.

وللتأكد من صدق التجربة في الغرض الذي صممت من أجله،وإمكانية إعطاء التجربة نفس النتائج إذا ما أعيد تطبيقها مرة أخرى تم استخدام اختبار معامل الثبات ألفا كرونباخ Alpha cronbach لقياس ثبات وصدق المحتوى لمتغيرات الدراسة. ويوضح الجدول رقم (2-3) نتائج هذا الاختبار والذي يتبين منه أن معامل الثبات لإجمالي درجات تأثير ما وراء البيانات على كفاءة استرجاع مقالات الدوريات الالكترونية العربية من محركات البحث عند البحث "بالعنوان"، بلغ (0.747) بما يدل على الثبات المرتفع الذي انعكس أثره على الصدق الذاتي (الذي يمثل الجذر التربيعي لمعامل الثبات) والذي بلغ (0.864).

أما معامل الثبات لإجمالي درجات تأثير ما وراء البيانات على كفاءة استرجاع مقالات الدوريات الالكترونية عند البحث "بالكلمات المفتاحية" فكان (0.776)، ما يدل على

الثبات المرتفع الذي انعكس أثره على الصدق الذاتي (الذي يمثل الجذر التربيعي لمعامل الثبات) والمقدر بـ (0.880).

الجدول رقم (3-2) معامل الثبات والصدق الذاتي لمتغيرات تأثير ما وراء البيانات على كفاءة استرجاع مقالات الدوريات الالكترونية من محركات البحث

المتغيرات	معامل ثبات ألفا كرونباخ	معامل الصدق
إجمالي: درجات تأثير ما وراء البيانات على كفاءة استرجاع مقالات الدوريات الالكترونية عند البحث بالعنوان	0.747	0.864
إجمالي: درجات تأثير ما وراء البيانات على كفاءة استرجاع مقالات الدوريات الالكترونية عند البحث بالكلمات المفتاحية	0.776	0.880

وفيما يلي عرضا لأبرز ما توصلت له الدراسة في كل محرك من محركات البحث الثلاثة.

1-2-1 تأثير ما وراء البيانات على كفاءة الاسترجاع من محرك البحث جوجل Google:

(أ) نتائج البحث بالعنوان:

يوضح الجدول رقم (3-3) مقالات مجموعات التجربة وفقا لترتيب ظهورها في قائمة نتائج محرك البحث جوجل عند البحث عن كل مقالة بعنوانها. مع ملاحظة أنه تم استخدام مقياس من (1) إلى (3)، وعند البحث عن أحد المقالات تم إعطاء المجموعة التي تظهر أولا في قائمة نتائج محركات البحث أعلى درجة على هذا المقياس أي (3)، والتي تظهر ثانيا (2)، والتي تظهر أخيرا (1). ليتم بعد ذلك تجميع الدرجات التي حققتها جميع مقالات كل مجموعة على حدة.

الجدول رقم (3-3) مقالات مجموعات التجربة وفقا لترتيب ظهورها
في قائمة نتائج البحث لجوجل عند البحث بالعنوان

مجموعات التجربة		أ	%	درجات المجموعة	ب	%	درجات المجموعة	ج	%	درجات المجموعة
رتبة الظهور في قائمة نتائج محرك البحث	أولا	35	%56.5	105	22	%35.5	66	9	%15	27
	ثانيا	19	%30.6	38	27	%43.5	54	16	%26.7	32
	ثالثا	8	%12.9	8	13	%21	13	35	%58.3	35
الإجمالي		62	%100	151	62	%100	133	60	%100	94
المتوسط الحسابي		2.44			2.15			1.57		
الانحراف المعياري		0.71			0.74			0.75		

يتبين من الجدول السابق ما يلي:

- استطاع محرك البحث جوجل استدعاء نسبة كبيرة من المقالات في مجموعات التجربة الثلاثة بما يدل على قدرته العالية على التكشيف؛ حيث بلغ أقصى عدد من المقالات التي استطاع استرجاعها لمجموعة من مجموعات التجربة (62) مقالة أي ما نسبته 86.11% من إجمالي مقالات المجموعة (72 مقالة)، وذلك لكل من المجموعة الضابطة (أ)، والمجموعة التجريبية (ب)، فيما استطاع محرك البحث جوجل استدعاء 60 مقالة للمجموعة التجريبية (ج) بما يعادل 83.33% من مقالات هذه المجموعة.
- عند المقارنة بين كل مجموعة من المجموعات وفقا لعدد مقالاتها التي ظهرت في المرتبة الأولى في قائمة نتائج محرك البحث، يتبين أن المجموعة الضابطة (أ) التي لا تحتوي مقالاتها على ما وراء البيانات تحتل مركز الصدارة بين مجموعات التجربة حيث بلغ عدد مقالاتها التي ظهرت أولا في قائمة نتائج محرك البحث (35) مقالة بنسبة 56.5% من إجمالي المقالات التي تم العثور عليها لهذه المجموعة، وجاءت في المرتبة الثانية

المجموعة التجريبية (ب) التي تحتوي على ما وراء البيانات وفقا لبؤرة دبلن، حيث بلغ عدد مقالاتها التي ظهرت أولا في قائمة نتائج البحث (22) مقالة بنسبة حوالي 35.5%. وفي المرتبة الأخيرة جاءت المجموعة التجريبية (ج) والتي ظهرت (9) مقالات منها أولا في قائمة نتائج البحث أي ما يعادل 15%.

- لتحديد اتجاه مقالات كل مجموعة نحو الظهور أولا أو ثانيا أو ثالثا في قائمة نتائج محرك البحث، تم مقارنة المتوسط الحسابي لكل مجموعة بفئات المتوسط المرجح وفقاً لمعايير الدرجة المرتفعة والمنخفضة، في إطار مقياس ليكرت الثلاثي الاتجاه Likert Scale المستخدم في هذه الدراسة وفقا لما يلي:

فئة المتوسط	الاتجاه
1.66-1.00	تميل مقالات المجموعة إلى (الظهور في المرتبة الثالثة والأخيرة في قائمة نتائج محرك البحث)
2.37-1.67	تميل مقالات المجموعة إلى (الظهور في المرتبة الثانية في قائمة نتائج محرك البحث)
3-2.38	تميل مقالات المجموعة إلى (الظهور في المرتبة الأولى في قائمة نتائج محرك البحث)

ووفقا لذلك فقد اتجهت مقالات المجموعة (أ) اتجاها عاما نحو الظهور في المرتبة الأولى في قائمة نتائج محرك البحث جوجل، وذلك بمتوسط قدره (2.44) وانحراف معياري (0.71). فيما حققت المجموعة (ب) اتجاها عاما نحو الظهور في المرتبة الثانية بمتوسط (2.15) وانحراف معياري (0.74). أما المجموعة (ج) فقد اتجهت نحو الظهور في المرتبة الأخيرة من بين مجموعات التجربة بمتوسط (1.57) وانحراف معياري (0.75).

الجدول رقم (3-4) نتائج اختبار كروسكال ويلز لمجموعات التجربة
عند البحث بالعنوان في جوجل

مجموعات التجربة	العدد	المتوسط الرتبي	قيمة كروسكال (كا^2)	القــرار	
				مستوى المعنوية	الدلالة
مجموعة أ	62	116.02	35.512	0.01	دالة
مجموعة ب	62	97.84			
مجموعة ج	60	62.68			

- تم وضع الفرض الصفري: "لا توجد فروق ذات دلالة إحصائية بين مجموعات التجربة عند البحث بالعنوان في جوجل". وتشير نتائج اختبار كروسكال ويلز الواردة في الجدول رقم (3-4) إلى وجود فروق ذات دلالة إحصائية بين مجموعات التجربة، حيث بلغت قيمة كروسكال ويلز (كا^2) (35.512)، وكان مستوى المعنوية أقل من (0.05) وهذا يعني رفض الفرض الصفري، وقبول وجود فروق ذات دلالة إحصائية بين مجموعات التجربة، إلا أن هذه الفروق أو التباين كانت لصالح المجموعة (أ) التي لا تحتوي على ما وراء البيانات والتي حققت أعلى متوسط رتبي (116.02) من بين مجموعات التجربة. ووفقا لهذه النتيجة يتضح عدم تأثير ما وراء البيانات على محرك البحث جوجل بل كانت الأفضلية للصفحات التي لا تشتمل على ما وراء البيانات.
- وفقا للنتيجة السابقة التي أشارت إلى تفوق المجموعة الضابطة (أ) التي لا تحتوي على ما وراء البيانات، يمكن قبول فرض الدراسة بالنسبة لمحرك البحث جوجل، والذي ينص على أنه "لا تؤثر ما وراء البيانات على قدرة محركات البحث على استرجاع مقالات الدوريات الإلكترونية العربية في مجال المكتبات والمعلومات". وتتفق هذه النتيجة مع دراسة Safari والتي أشارت إلى عدم تأثير ما وراء البيانات على أداء محرك البحث جوجل[1]، وكذلك مع ما صرح به جوجل من عدم استخدام تيجان ما وراء البيانات

(1) Safari, Mehdi. Search Engines and Resource Discovery on the Web: Is Dublin Core an Impact Factor? [online]. Webology. Vol. 2, No. 2 (August 2005). Available at: http://www.webology.org/2005/v2n2/a13.html. [Accessed 20/1/2016].

في ترتيب الصفحات Ranking في صفحة نتائج البحث، ولكنه قد يقوم بالاستعانة بها في طريقة عرض المواقع في صحفة نتائج محرك البحث وتوفير المقتطف الغني Rich Snippet[1].

ب- نتائج البحث بالكلمات المفتاحية:

تم البحث عن كل مقالة من المقالات باستخدام الكلمات المفتاحية التي تم تضمينها في ما وراء بيانات المقالات، حيث تم إعداد استراتيجية تتناسب مع كل مقالة، وعندما يتطلب الأمر البحث بأكثر من كلمة مفتاحية تم الربط بينهما بعلامة الزائد، ويعرض الجدول رقم (3-5) مقالات كل مجموعة من مجموعات التجربة وفقا لترتيب ظهورها في نتائج البحث لجوجل عند البحث بالكلمات المفتاحية.

الجدول رقم (3-5) مقالات مجموعات التجربة وفقا لترتيب ظهورها في قائمة نتائج البحث لجوجل عند البحث بالكلمات المفتاحية

مجموعات التجربة		أ	%	درجات المجموعة	ب	%	درجات المجموعة	ج	%	درجات المجموعة
رتبة الظهور في قائمة نتائج محرك البحث	أولا	5	62.5	15	5	71.4	15	1	20	3
	ثانيا	2	25	4	1	14.3	2	2	40	4
	ثالثا	1	12.5	1	1	14.3	1	2	40	2
الإجمالي		8	100	20	7	100	18	5	100	9
المتوسط الحسابي		2.50			2.57			1.80		
الانحراف المعياري		0.75			0.78			0.83		

(1) Meta tags - Custom Search Help [online]. Available at: https://support.google.com/customsearch/answer/2595557?hl=en. [Accessed 27/5/2014]

- يعد معدل الاستدعاء الذي حققه جوجل عند البحث بالكلمات المفتاحية ضعيف مقارنة بمعدل الاستدعاء الذي حققه من قبل عند البحث بالعنوان، وفي هذا إشارة إلى اعتماد جوجل بشكل أكبر على تكشيف النص الكامل عند الاسترجاع وليس ما ورد في عناصر ما وراء البيانات وتحديدا تاج الكلمات المفتاحية. فقد بلغ أقصى عدد من المقالات التي تم استرجاعها (8) مقالات بنسبة 11.11% من إجمالي مقالات المجموعة، وقد انتمت هذه المقالات إلى المجموعة (أ). بينما كان أقل عدد من المقالات التي تم استرجاعها في مجموعة واحدة هو (5) مقالات أي ما يعادل 6.94%، وتخص هذه المقالات المجموعة التجريبية (ج). ويمكن تفسير ذلك بأن محرك البحث جوجل يعتمد في تقديم نتائج البحث وفقا لاستفسار معين أو أحد الكلمات المفتاحية على ما سبق أن اختزنه وسجله من مواقع قام المستفيدون بالضغط عليها واختيارها عند البحث بهذه الاستفسارات أو الكلمات المفتاحية، يؤكد ذلك ظهور مقالات الدوريات عينة التجربة ولكن على المواقع الأصلية للدوريات التي تتاح عليها.
- اتجهت مقالات المجموعة الضابطة (أ) نحو الظهور في المرتبة الأولى بمتوسط حسابي يقدر ب (2.50) وانحراف معياري (0.75) وذلك عند مقارنتها بفئات المتوسط السابق تحديدها. وكذلك المجموعة (ب) التي تحتوي على ما وراء البيانات وفقا لبؤرة دبلن بمتوسط بلغ (2.57) وانحراف معياري (0.78). بينما اتجهت مقالات المجموعة (ج) التي تحتوي على ما وراء البيانات وفقا لمفردات سكيما إلى الظهور في المرتبة الثانية بمتوسط (1.80).
- بلغ عدد مقالات المجموعة (أ) التي ظهرت أولا في قائمة نتائج محرك البحث جوجل عند البحث بالكلمات المفتاحية (5) مقالات أي ما نسبته 62.5% من مقالات المجموعة التي استرجعها محرك البحث جوجل. وفي نفس المرتبة ظهر نفس العدد من المقالات (5 مقالات) من المجموعة التجريبية (ب) ولكن بنسبة تقدر بنحو 71.4% من مقالات هذه المجموعة التي استرجعها محرك البحث جوجل. بينما جاءت في المرتبة الأخيرة المجموعة التجريبية (ج) حيث ظهرت مقالة واحدة من هذه المجموعة في المرتبة الأولى في قائمة نتائج البحث أي ما يعادل 20%.

الجدول رقم (3-6) نتائج اختبار كروسكال ويلز لمجموعات التجربة
عند البحث بالكلمات المفتاحية على جوجل

عينة الدراسة	العدد	المتوسط الرتبي	قيمة كروسكال ($كا^2$)	القرار	
				مستوى المعنوية	الدلالة
مجموعة أ	8	11.44	3.264	0.19	غير دالة
مجموعة ب	7	12.07			
مجموعة ج	5	6.80			

- تم وضع الفرض الصفري: "لا توجد فروق ذات دلالة إحصائية بين مجموعات التجربة عند البحث بالكلمات المفتاحية في جوجل". ويتبين من الجدول السابق (3-6) عدم وجود فروق ذات دلالة إحصائية بين مجموعات التجربة، حيث بلغت قيمة كروسكال ويلز ($كا^2$) (3.264) وبمستوى معنوية 0.19 وهذا المستوى أكبر من 0.05 وهذا يعني قبول الفرض الصفري بعدم وجود فروق ذات دلالة إحصائية بين المجموعات. وبالتالي يمكن قبول فرض الدراسة الذي ينص على أن ما وراء البيانات لا تؤثر على كفاءة محرك البحث جوجل في استرجاع مقالات الدوريات الإلكترونية العربية في مجال المكتبات والمعلومات.

2-2-1 تأثير ما وراء البيانات على كفاءة استرجاع المقالات من محرك البحث ياهو Yahoo:

(أ) نتائج البحث بالعنوان:

تجدر الإشارة إلى ملاحظة هامة هي أن الباحثة قامت باستخدام محرك البحث ياهو Yahoo search وليس محرك البحث ياهو مكتوب Maktoob، حيث لوحظ اختلاف نتائج البحث على كل منهما.

الجدول رقم (3-7) مقالات مجموعات التجربة وفقا لترتيب ظهورها
في قائمة نتائج البحث لياهو عند البحث بالعنوان

مجموعات التجربة		أ	%	درجات المجموعة	ب	%	درجات المجموعة	ج	%	درجات المجموعة
رتبة الظهور في قائمة نتائج محرك البحث	أولا	6	85.7	18	7	100	21	5	100	15
	ثانيا	1	14.3	2	-	-	-	-	-	-
	ثالثا	-	-	-	-	-	-	-	-	-
الإجمالي		7	100	20	7	100	21	5	100	15
المتوسط الحسابي		2.86			3.0			3.0		
الانحراف المعياري		0.37			0			0		

يتبين من الجدول رقم (3-7) ما يلي:

- يعد معدل الاستدعاء الذي حققه محرك البحث ياهو لمقالات مجموعات التجربة عند البحث بالعنوان ضعيف، حيث بلغ أقصى عدد من المقالات التي تم استرجاعها لأحد مجموعات التجربة سبعة (7) مقالات بنسبة تقدر بنحو 9.72% من إجمالي مقالات المجموعة، وذلك لكل من المجموعة الضابطة (أ)، والمجموعة التجريبية (ب). بينما كان عدد المقالات التي تم استرجاعها من مقالات المجموعة التجريبية (ج) هو خمسة مقالات (5) بنسبة حوالي 6.94% من إجمالي مقالات المجموعة.
- وفقا لما ورد في الجدول رقم (3-7)، يتبين أن أقصى عدد من مقالات المجموعة الأولى (6 مقالات) ظهر أولا في قائمة نتائج محرك البحث ياهو بما يعادل 85.7% من إجمالي ما استرجعه هذا المحرك من مقالات هذه المجموعة، بينما بلغت نسبة المقالات التي ظهرت أولا للمجموعة (ب) 100% بعدد سبع مقالات (7). وحققت نفس النسبة المجموعة (ج) 100% حيث ظهر لهذه المجموعة خمس مقالات فقط ظهرت جميعها في المرتبة الأولى في قائمة نتائج محرك البحث ياهو. ويؤكد ذلك متوسطات المجموعات التي أشارت إلى اتجاه مقالات مجموعات التجربة الثلاثة نحو الظهور في المرتبة الأولى

في قائمة محرك البحث ياهو حيث كانت 2.86 للمجموعة (أ)، و3.0 لكل من المجموعتين التجريبيتين (ب)، و(ج).

الجدول رقم (3-8) نتائج اختبار كروسكال ويلز لمجموعات التجربة عند البحث بالعنوان على ياهو

عينة الدراسة	العدد	المتوسط الرتبي	قيمة كروسكال (كا^2)	القـرار	
				مستوى المعنوية	الدلالة
مجموعة أ	7	9.14	1.714	0.42	غير دالة
مجموعة ب	7	10.50			
مجموعة ج	5	10.50			

- تم اختبار الفرض الصفري: "لا توجد فروق ذات دلالة إحصائية بين مجموعات التجربة عند البحث بالعنوان في ياهو". ويظهر من الجدول (3-8) عدم وجود فروق ذات دلالة إحصائية بين كفاءة استرجاع مجموعات التجربة من محرك البحث ياهو عند البحث بالعنوان، حيث بلغت قيمة كروسكال ويلز (كا^2) (1.714)، وذلك عند مستوى معنوية أكبر من (0.05) وهذا يعني قبول الفرض الصفري. وبناء على هذه النتيجة يمكن قبول فرض الدراسة بأن ما وراء البيانات لا تؤثر على كفاءة استرجاع محرك البحث ياهو لمقالات الدوريات الإلكترونية العربية المتخصصة في المجال.
- تجدر الإشارة إلى ملاحظة هامة أنه على الرغم من تبين عدم تأثير ما وراء البيانات على كفاءة الاسترجاع من محرك البحث ياهو إلا أنه عند عرض هذا المحرك لعدد من مقالات المجموعة التجريبية (ب) في قائمة النتائج، تم استخدام الملخص الذي قامت الباحثة بإدراجه في عنصر الوصف لبؤرة دبلن DCTERMS.Description أسفل رابط الموقع كما في الشكل رقم (3-5)، والذي يختلف عن النص أو الملخص الوارد في المقالة بما يؤكد على فهم واستخدام محرك البحث ياهو لما وراء البيانات وفقا لبؤرة دبلن وخاصة في عرض هذه المقالات في قائمة نتائج البحث، وهذا قد يؤدي في حد ذاته إلى زيادة معدل زيارات هذه المقالات.

"النشر الالكتروني ومستقبل الكلمة المطبوعة" - Yahoo Search Results

Mail Search News Sports Finance Celebrity Weather Answers Flickr Mobile

"النشر الالكتروني ومستقبل الكلمة المطبوعة" Search

Web Images Video Anytime

النشر الإلكتروني ومستقبل الكلمة المطبوعة
www.ajainlis.com/elarbia-3000/النشر...ومستقبل...
يتناول البحث بدايات ظهور الكتاب الإلكتروني والاهتمام بتصنيع الأقراص المكتنزة ومستقبل ...

<meta name="DCTERMS.Description" content="يتناول البحث بدايات ظهور الكتاب الإلكتروني والاهتمام بتصنيع الأقراص المكتنزة ومستقبل الكتاب الورقي والكلمة المطبوعة في ظل النشر الإلكتروني">

الشكل رقم (3-5) استخدام ياهو لعنصر الوصف وفقا لبؤرة دبلن
عند عرض الصفحات في قائمة نتائج البحث

(ب) نتائج البحث بالكلمات المفتاحية:

يوضح الجدول التالي رقم (3-9) ترتيب ظهور مقالات كل مجموعة في قائمة نتائج محرك البحث ياهو عند البحث عن كل مقالة بالكلمات المفتاحية، والدرجات التي حققتها كل مجموعة على مقياس الدراسة من 3 إلى 1.

يتضح من الجدول (3-9) النقاط التالية:

- حقق محرك البحث ياهو معدل استدعاء مرتفع لمقالات مجموعات التجربة مقارنة بما حققه عند البحث بالعنوان، حيث كان أكبر عدد من المقالات التي تم استرجاعها هو (18) مقالة بنسبة تقدر بنحو 25% من إجمالي مقالات المجموعة، وقد انتمت هذه المقالات إلى المجموعة التجريبية (ب)، بينما تم استرجاع (15) مقالة للمجموعة التجريبية (ج) أي ما يوازي 20.83%. بينما كان أقل معدل للاستدعاء هو (11) مقالة بما يعادل 15.28% وذلك لمقالات المجموعة التجريبية (أ).

الجدول رقم (3-9) مقالات مجموعات التجربة وفقا لترتيب ظهورها
في قائمة نتائج البحث لياهو عند البحث بالكلمات المفتاحية

مجموعات التجربة		أ	%	درجات المجموعة	ب	%	درجات المجموعة	ج	%	درجات المجموعة
رتبة الظهور في قائمة نتائج محرك البحث	أولا	10	90.9	30	17	94.4	51	13	86.7	39
	ثانيا	1	9.1	2	1	5.6	2	1	6.7	2
	ثالثا	-	-	-	-	-	-	1	6.7	1
الإجمالي		11	100	32	18	100	53	15	100	42
المتوسط الحسابي		2.91			2.94			2.80		
الانحراف المعياري		0.30			0.23			0.56		

- تتقارب مجموعات التجربة من حيث المقالات التي ظهرت لها في المرتبة الأولى في قائمة نتائج بحث ياهو، حيث بلغت نسبة هذه المقالات في المجموعة الضابطة (أ) 90.9% (10 مقالات)، وبنسبة أعلى قليلا جاءت المجموعة التجريبية (ب) حيث بلغت نسبة مقالاتها التي ظهرت أولا في قائمة النتائج لهذا المحرك 94.4%، بينما بلغت نسبة مقالات المجموعة التجريبية (ج) التي ظهرت أولا في قائمة نتائج البحث 86.7% بعدد (13) مقالة.
- حققت مقالات مجموعات التجربة الثلاثة عند البحث في ياهو باستخدام الكلمات المفتاحية اتجاها نحو الظهور في المرتبة الأولى، حيث بلغ المتوسط في المجموعة (أ) (2.91) وبانحراف معياري (0.30). وكان المتوسط للمجموعة التجريبية (ب) (2.94) وانحراف قدره (0.23). وبلغ المتوسط للمجموعة التجريبية (ج) (2.80)، والانحراف المعياري (0.56).

الجدول رقم (3-10) نتائج اختبار كروسكال ويلز لمجموعات التجربة
عند البحث بالكلمات المفتاحية على ياهو

عينة الدراسة	العدد	المتوسط الرتبي	قيمة كروسكال (كا^2)	القــرار	
				مستوى المعنوية	الدلالة
مجموعة أ	11	22.55	0.651	0.72	غير دالة
مجموعة ب	18	23.31			
مجموعة ج	15	21.50			

- تم وضع الفرض الصفري: "لا توجد فروق ذات دلالة إحصائية بين مجموعات التجربة عند البحث بالكلمات المفتاحية في ياهو". ويوضح الجدول رقم (10/3) نتائج اختبار كروسكال ويلز لمجموعات التجربة، والذي أشار إلى عدم وجود فروق ذات دلالة إحصائية بين مجموعات التجربة، حيث بلغت قيمة كروسكال ويلز (كا^2) (0.651)، وذلك عند مستوى معنوية أكبر من (0.05).
- تؤكد هذه النتيجة على ما تم التوصل له سابقا من عدم وجود فروق ذات دلالة إحصائية بين مجموعات التجربة عند البحث بالعنوان في ياهو، ووفقا لذلك يمكن قبول فرض الدراسة الذي ينص على: "لا تؤثر ما وراء البيانات على قدرة محركات البحث – وتحديدا محرك البحث ياهو - على استرجاع مقالات الدوريات الإلكترونية العربية في مجال المكتبات والمعلومات". إلا أن هذه النتيجة حول تأثير ما وراء البيانات على كفاءة الاسترجاع من محرك البحث ياهو تتعارض مع ما توصلت له دراسة زهانج[1] من تأثير ما وراء البيانات وبشكل خاص المكودة وفقا لبؤرة دبلن على عدد من محركات البحث من بينها محرك البحث ياهو.

(1) Zhang, Jin and Dimitroff, Alexandra. Internet Search Engines' Response to Metadata Dublin Core Implementation. Journal of Information Science. Vol. 30, No.4 (2004). P. 315.

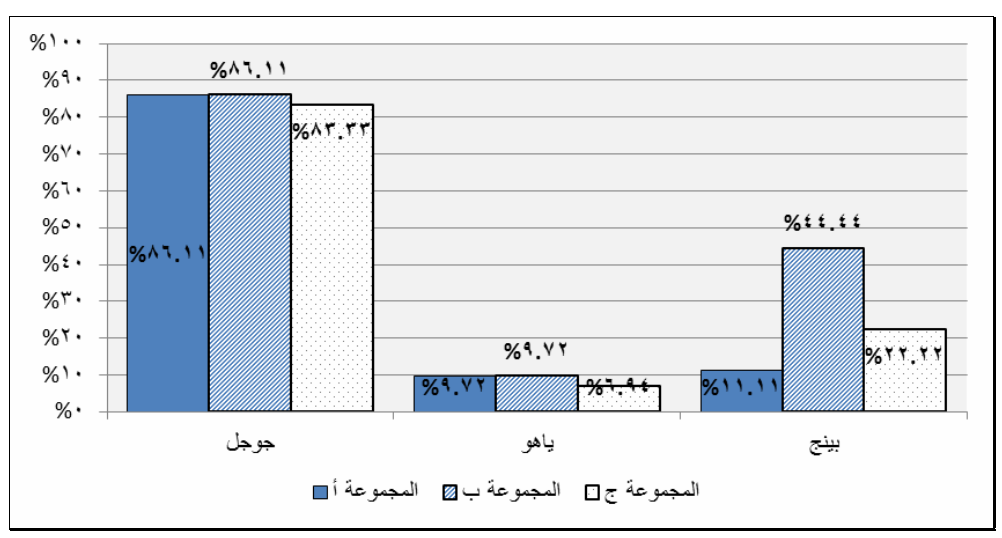
3-2-1 تأثير ما وراء البيانات على كفاءة استرجاع المقالات من محرك البحث بينج Bing:

(أ) نتائج البحث بالعنوان:

الجدول رقم (3-11) مقالات مجموعات التجربة وفقا لترتيب ظهورها
في قائمة نتائج البحث لبينج عند البحث بالعنوان

مجموعات التجربة		أ	%	درجات المجموعة	ب	%	درجات المجموعة	ج	%	درجات المجموعة
رتبة الظهور في قائمة نتائج محرك البحث	أولا	7	87.5	21	31	96.9	93	13	81.2	39
	ثانيا	1	12.5	2	1	3.1	2	3	18.8	6
	ثالثا	-	-		-	-	-	-	-	-
الإجمالي		8	100	23	32	100	95	16	100	45
المتوسط الحسابي		2.88			2.97			2.81		
الانحراف المعياري		0.35			0.17			0.40		

- بلغ أقصى عدد من المقالات التي استطاع محرك البحث بينج استدعائها لأحد مجموعات التجربة عند البحث بالعنوان (32) مقالة بنسبة 44.44% كانت جميعها مقالات للمجموعة التجريبية (ب)، محققا بذلك معدل استدعاء متوسط من بين محركات البحث محل الدراسة كما يتضح من الشكل رقم (3-6)، وجاءت بعد ذلك مقالات المجموعة التجريبية (ج) حيث تم استرجاع (16) مقالة لهذه المجموعة بنسبة بلغت نحو 22.22%. لتأتي في المرتبة الأخيرة المجموعة الضابطة (أ) والتي تم استرجاع (8) مقالات من هذه المجموعة أي ما نسبته 11.11%.



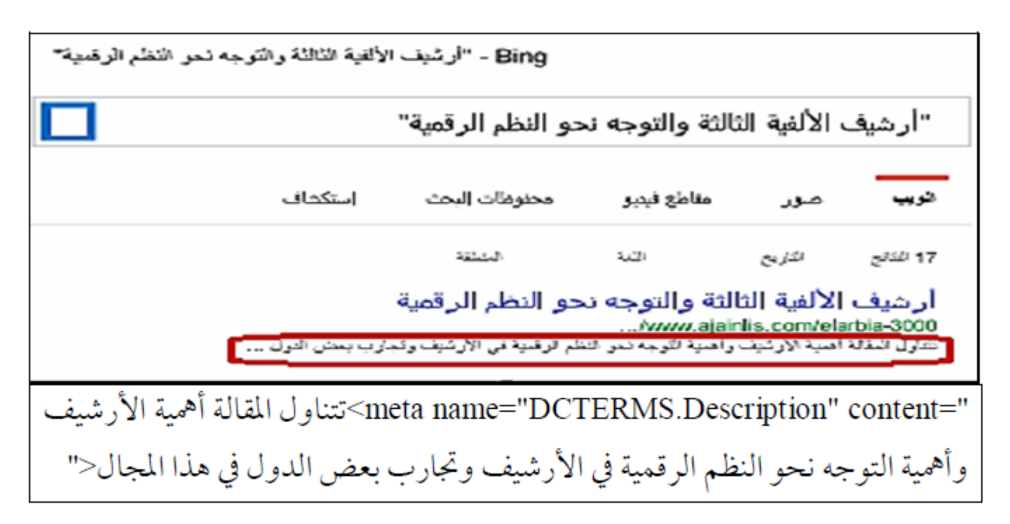
الشكل رقم (3-6) معدل استدعاء مقالات مجموعات التجربة من محركات البحث محل الدراسة عند البحث بالعنوان

- عند المقارنة بين مجموعات التجربة من حيث نسبة ظهور مقالاتها في المرتبة الأولى يتبين أن المجموعة التجريبية (ب) حققت أعلى نسبة مقالات ظهرت أولا في قائمة محرك البحث بينج حيث بلغ عدد هذه المقالات (31) مقالة بما يعادل نسبة حوالي 96.9% من إجمالي المقالات المسترجعة لهذه المجموعة، تلتها المجموعة الضابطة (أ) حيث بلغت نسبة المقالات التي وردت أولا في قائمة محرك البحث 87.5%، وجاءت بعد ذلك المجموعة التجريبية (ج) حيث كانت نسبة مقالاتها التي ظهرت أولا في قائمة النتائج 81.2% من إجمالي مقالات هذه المجموعة التي استرجعها بينج.
- بمقارنة المتوسطات الحسابية التي حققتها كل مجموعة من مجموعات التجربة بفئات المتوسطات السابق تحديدها في إطار المقياس الثلاثي المستخدم في هذه الدراسة، يتبين أن مجموعات التجربة الثلاثة حققت اتجاها عاما نحو ظهور مقالاتها في المرتبة الأولى حيث كان المتوسط لكل من المجموعة (أ، ب، ج) على التوالي (2.88، 2.97، و2.81).

الجدول رقم (3-12) نتائج اختبار كروسكال ويلز لمجموعات التجربة
عند البحث بالعنوان على بينج

عينة الدراسة	العدد	المتوسط الرتبي	قيمة كروسكال ($كا^2$)	القـرار	
				مستوى المعنوية	الدلالة
مجموعة أ	8	27.50	3.289	0.19	غير دالة
مجموعة ب	32	30.13			
مجموعة ج	16	25.75			

- تم وضع الفرض الصفري: "لا توجد فروق ذات دلالة إحصائية بين مجموعات التجربة عند البحث بالعنوان في بينج". وتشير البيانات الواردة في الجدول رقم (3-12) إلى قبول هذا الفرض حيث بلغت قيمة كروسكال ويلز ($كا^2$) (3.289)، وذلك عند مستوى معنوية أكبر من (0.05). وبناء على ذلك يمكن قبول فرض الدراسة الذي يشير إلى عدم تأثير ما وراء البيانات على محركات البحث – وتحديدا محرك البحث بينج- عند استرجاع مقالات الدوريات الإلكترونية العربية في مجال المكتبات.
- لوحظ استخدام محرك البحث بينج للملخص الذي يرد في تاج الوصف لبؤرة دبلن DCTERMS. Description كما يظهر في الشكل رقم (3-7)، وفي هذا دلالة على فهم واستخدام محرك البحث بينج لما وراء البيانات المكودة وفقا لبؤرة دبلن في عرض المواقع في قائمة نتائج محرك البحث بينج، وهذا يتفق مع استخدام محرك البحث ياهو لنفس التاج كما ورد سابقا.



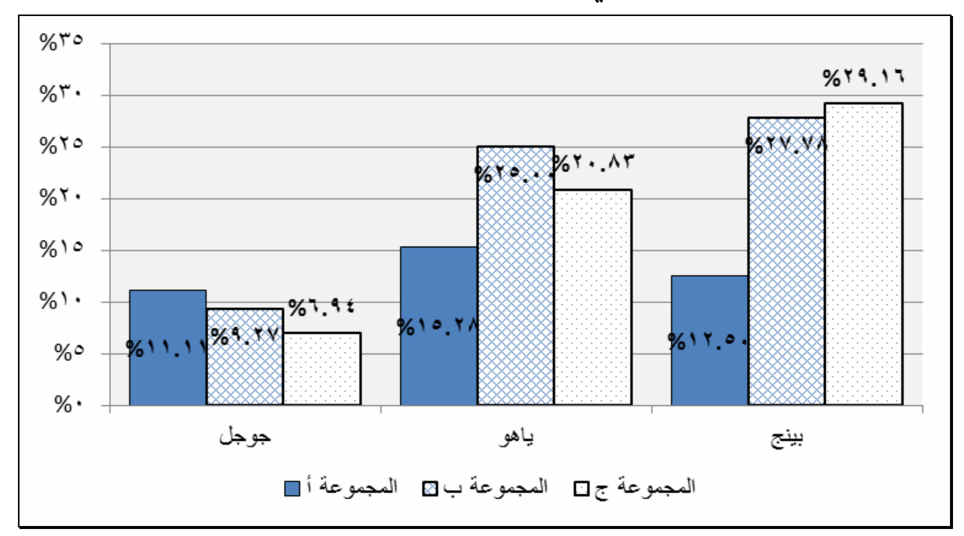
الشكل رقم (3-7) استخدام بينج لعنصر الوصف وفقا لبؤرة دبلن عند عرض الصفحات في قائمة نتائج البحث

(ب) نتائج البحث بالكلمات المفتاحية:

الجدول رقم (3-13) مقالات مجموعات التجربة وفقا لترتيب ظهورها في قائمة نتائج البحث لبينج عند البحث بالكلمات المفتاحية

مجموعات التجربة		أ	%	درجات المجموعة	ب	%	درجات المجموعة	ج	%	درجات المجموعة
رتبة الظهور في قائمة نتائج محرك البحث	أولا	9	100	27	18	90	54	20	95.2	60
	ثانيا	-	-	-	2	10	4	1	4.8	2
	ثالثا	-	-	-	-	-	-	-	-	-
الإجمالي		9	100	27	20	100	58	21	100	62
المتوسط الحسابي		3.0			2.90			2.95		
الانحراف المعياري		0			0.30			0.21		

- تفوق محرك البحث بينج على غيره من محركات البحث عينة الدراسة في معدل الاستدعاء عند البحث بالكلمات المفتاحية كما يتضح من الشكل رقم (3-8)؛ حيث بلغ أقصى عدد للمقالات التي استرجعها لمجموعة من مجموعات التجربة (21) مقالة تنتمي للمجموعة التجريبية (ج) وبنسبة حوالي 29.16% بينما حقق ياهو وجوجل معدلات استدعاء أقل عند البحث بالكلمات المفتاحية كما ورد سابقا (25%، و11.11% على التوالي). كما استطاع بينج استدعاء (20) مقالة من مقالات المجموعة التجريبية (ب) أي ما يوازي نحو 27.78% من مقالات هذه المجموعة. وجاءت في المرتبة الأخيرة مقالات المجموعة الضابطة (أ) حيث تم استرجاع عدد تسعة (9) مقالات فقط من مقالاتها بنسبة بلغت 12.5% من إجمالي مقالات هذه المجموعة.



الشكل رقم (3-8) معدل استدعاء مقالات مجموعات التجربة من محركات البحث محل الدراسة عند البحث بالكلمات المفتاحية

وردت جميع المقالات التي تم استرجاعها للمجموعة الضابطة (أ) في المرتبة الأولى (100%)، بينما ورد في نفس هذه المرتبة حوالي 95.2% من مقالات المجموعة التجريبية (ج)، وجاءت بعد ذلك المجموعة التجريبية (ب) والتي ظهر في المرتبة الأولى حوالي 90% من مقالاتها.

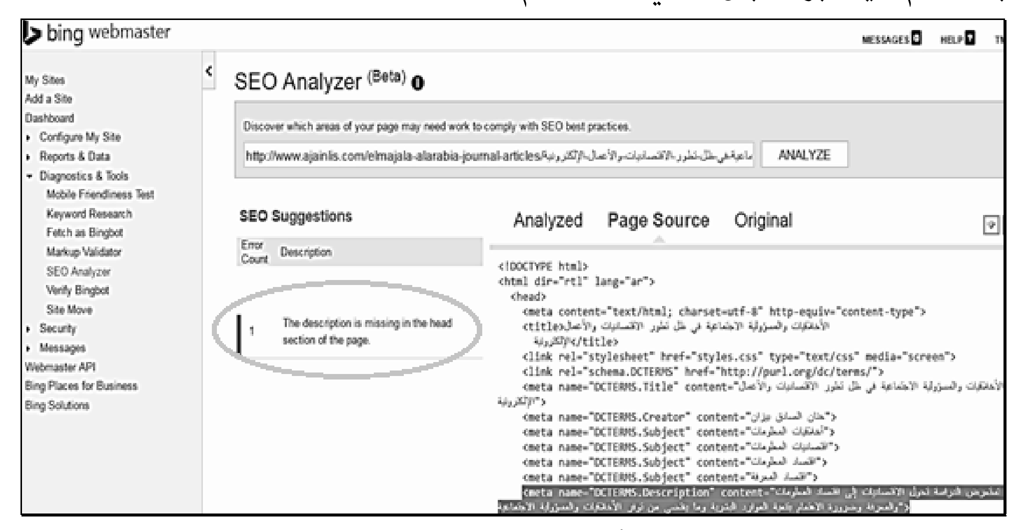
الجدول رقم (3-14) نتائج اختبار كروسكال ويلز لمجموعات التجربة عند البحث بالكلمات المفتاحية على بينج

عينة الدراسة	العدد	المتوسط الرتبي	قيمة كروسكال (كا^2)	القــرار	
				مستوى المعنوية	الدلالة
مجموعة أ	9	27.0	1.175	0.55	غير دالة
مجموعة ب	20	24.50			
مجموعة ج	21	25.81			

- تم وضع الفرض الصفري "لا توجد فروق ذات دلالة إحصائية بين مجموعات التجربة عند البحث بالكلمات المفتاحية على بينج". ووفقا لنتائج اختبار كروسكال كما ترد في الجدول رقم (3-14) فقد بلغت قيمة كروسكال (كا^2) (1.175)، وذلك عند مستوى معنوية (0.55) ونظرا لأن مستوى المعنوية أكبر من 0.05 فقد تم قبول الفرض الصفري. وبناء على ذلك يمكن قبول فرض الدراسة بأن ما وراء البيانات لا تؤثر على كفاءة محركات البحث في استرجاع مقالات الدوريات الإلكترونية العربية في المجال. وتتفق هذه النتيجة مع النتيجة السابقة التي تم التوصل إليها في نتائج البحث بالعنوان على محرك البحث بينج.

بناء على ما سبق، فإن نتائج الاختبارات الإحصائية لمحركات البحث الثلاثة (جوجل، وياهو، وبينج) تؤدي إلى قبول فرض الدراسة بعدم تأثير ما وراء البيانات على كفاءة الاسترجاع من محركات البحث. وعلى الرغم من ذلك فمن المهم الإشارة إلى أن هذه الدراسة عملت على دراسة ما وراء البيانات الواردة وفقا لخطط بعينها وهي بؤرة دبلن، ومفردات الاسكيما وفقا لإطار وصف المصدر RDFa وقد تكون ما وراء البيانات وفقا لهذه الخطط غير مؤثرة بالنسبة لمحركات البحث ولا تضعها في اعتبارها فيما يتعلق بكفاءة الاسترجاع بشكل خاص، ويمكن لما وراء البيانات إذا ما وردت وفقا لخطط أخرى مثل تيجان الميتا META HTML Tags أو باستخدام البيانات الدقيقة Microdata

فمن المحتمل أن يكون لها تأثير وهذا ما يحتاج إلى دراسة أخرى لتحديد مدى صحته، وقد يرجح هذا التصور توصية محركات البحث بالاهتمام بإدراج عنصر العنوان والوصف، وإعلان محرك البحث ياهو على أنه لا يزال يدعم تاج الكلمات المفتاحية إلا أنه يحظى بأقل قدر من الأهمية عند ترتيب النتائج(1). كما قامت الباحثة باختبار بسيط على أدوات مشرفي المواقع لمحرك البحث بينج حيث تم استخدام أداة تحليل تحسين أداء محركات البحث SEO Analyzer لأحد صفحات المجموعة التجريبية (ب) التي تحتوي على ما وراء البيانات وفقا لدبلن لتظهر رسالة بوجود خطأ حول عدم توافر عنصر الوصف على الرغم من وجوده باستخدام صيغة بؤرة دبلن كما في الشكل رقم (3-9).



الشكل رقم (3-9) نتيجة استخدام أداة SEO Analyzer لأحد مقالات المجموعة التجريبية (ب)

(1) Meta element [online]. In: Wikipedia: The Free Encyclopedia. Available at: https://en.wikipedia.org/wiki/Meta_element. [Accessed 11/9/2015].

2 إمكانيـة اسـتثمار مـا وراء البيانـات في مقـالات الـدوريات الإلكترونيـة العربيـة في مجـال المكتبات والمعلومات:

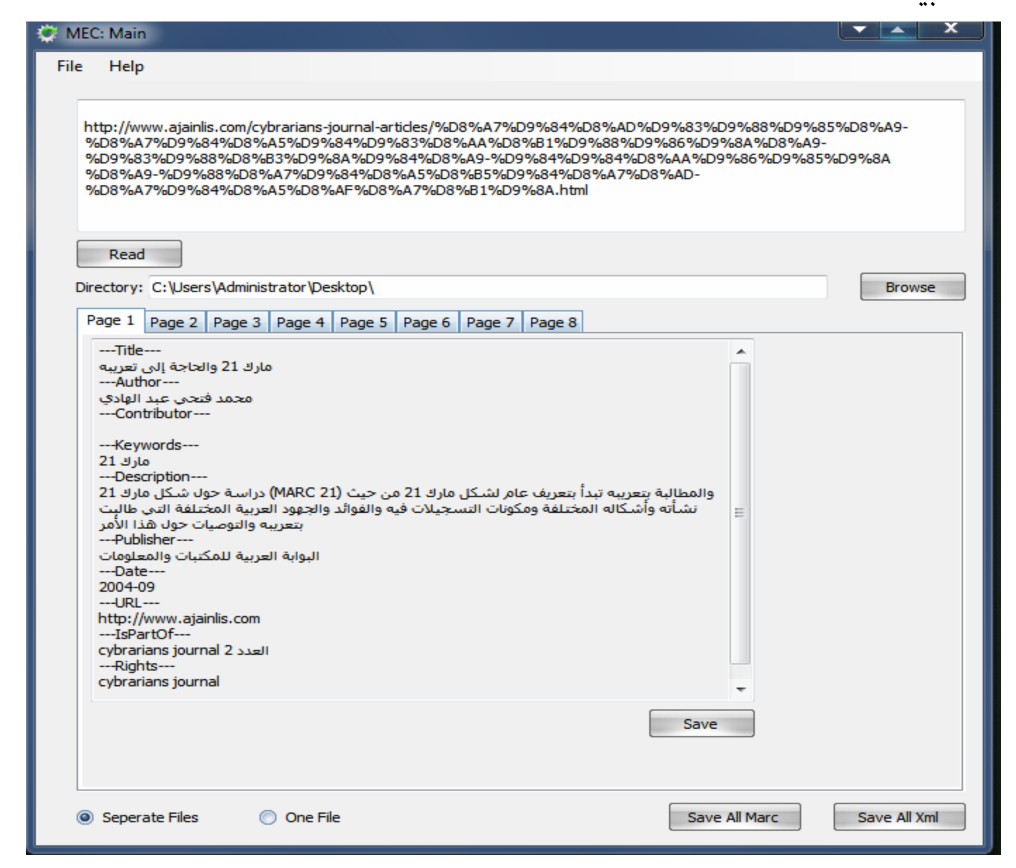
يتناول هذا الجزء مدى إمكانية استثمار ما وراء البيانات في صفحات العنكبوتيـة بصـفة عامة وفي مقالات الدوريات الإلكترونية العربية في مجال المكتبات والمعلومـات بصـفة خاصـة في إعداد تسجيلات ببليوجرافية وفقا لشكل الاتصـال مـارك 21 MARC21 يمكـن إيـداعها في فهارس المكتبات أو أحد المستودعات الرقمية، أو في إعداد كشاف لهذه المقالات.

ولاختبار هذه الإمكانية تم وضع فرض ينص على أنه " لا يمكن استثمار ما وراء البيانات في مقالات الدوريات الإلكترونية العربية في مجال المكتبـات والمعلومـات في إعـداد تسـجيلات ببليوجرافية وفقا لشكل الاتصال مارك 21 MARC21"، ولأجل اختبار هذا الفرض تـم القيـام بالخطوات التالية:

1- إعداد برنامج لاستخراج وتحويل ما وراء البيانات من الصفحات المتاحة على العنكبوتيـة وللقيام بإعداد هذا البرنامج تم الاستعانة بأحد مصممي البرامج، وقامت الباحثة بتحديد المواصفات المطلوبة لهذا البرنامج والتي تمثلت في أن يكون قادر على:
   - استخراج ما وراء البيانات من المقالات عينة التجربة سواء وردت وفقا لبـؤرة دبلـن، أو وفقا لمفردات سكيما Schema.org، وكذلك يكـون قـادرا عـلى اسـتخراج مـا وراء البيانات الـواردة في تيجـان الميتـا HTML Meta Tags، وذلـك عـن طريـق إدخـال الموقع الموحد للمصدر URL لكل مقالة من المقالات.
   - يستطيع حفظ ما وراء البيانات التي تم اسـتخراجها في ملـف XML، وكـذلك ملـف MARC/XML.

2- إدخال المواقع URL الخاصة بمقالات المجموعة التجريبية (ب) التي تشتمل على مـا وراء البيانات وفقا لبؤرة دبلن، وكذلك مواقع مقالات المجموعة التجريبية (ج) التـي تحتوي عـلى مـا وراء البيانـات وفقـا لمفـردات سـكيما في برنـامج اسـتخراج مـا وراء البيانات الذي تم إعداده. وقد استطاع البرنامج اسـتخلاص مـا وراء البيانـات بنجـاح

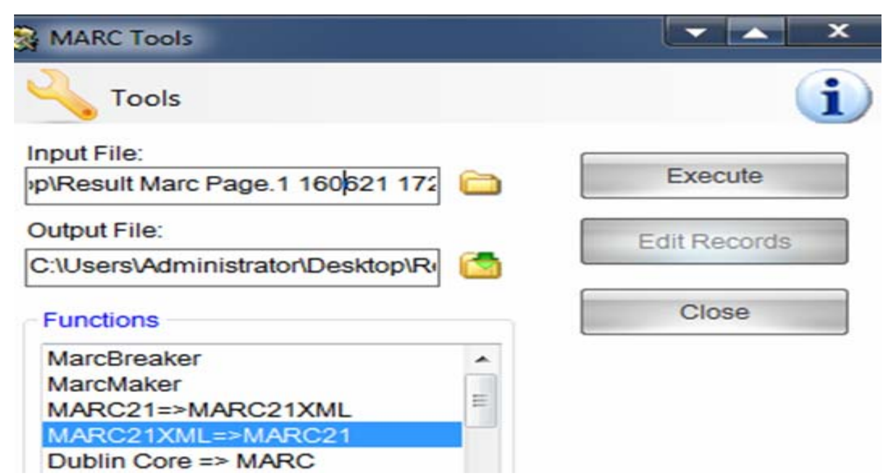
لعدد بلغ (65) مقالة لكل مجموعة أي ما نسبته 90.27% من إجمالي مقالات المجموعة الواحدة (72 مقالة). ويوضح الشكل التالي رقم (3-10) واجهة برنامج استخراج ما وراء البيانات.



الشكل رقم (3-10) واجهة برنامج استخراج وتحويل ما وراء البيانات الذي تم إعداده للدراسة

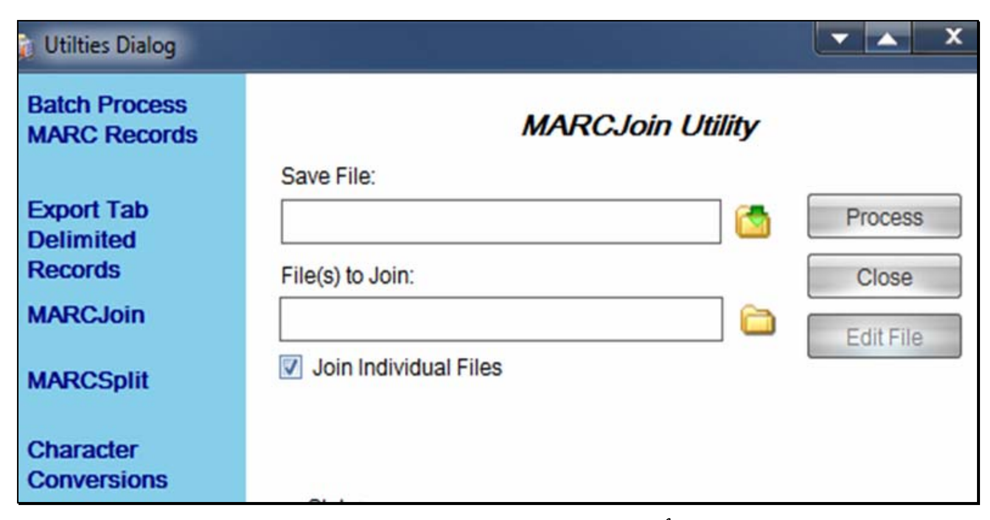
ومن الجدير بالذكر، أن هذا البرنامج لا يستطيع حفظ جميع تسجيلات ما وراء بيانات المقالات التي تم استخراجها في ملف واحد، مما يؤدي إلى التعامل مع كل ملف خاص بكل مقالة على حدة عند تحويله إلى صيغة مارك MARC في برنامج MARC Edit كما سيتضح فيما يلي، ويمكن أن يشكل ذلك أحد التعديلات على البرنامج فيما بعد.

3- الاستعانة ببرنامج MARC Edit، وهو برنامج مجاني متاح على الإنترنت، في تحويل ملفات ما وراء البيانات التي تم استخراجها بصيغة MARC/XML إلى ملفات بصيغة MARC21 عن طريق الخاصية MARC Tools كما يظهر في الشكل رقم (3-11)، واستطاع البرنامج بنجاح تحويل جميع الملفات التي تم استخراجها إلى ملفات بصيغة مارك .mrc



الشكل رقم (3-11) تحويل ملفات ما وراء البيانات التي تم استخراجها من صيغة MARC/XML إلى MARC باستخدام برنامج مارك إيدت MARC Edit

كما تمت الاستعانة بهذا البرنامج في عمل ملف واحد لجميع التسجيلات التي تم تحويلها لصيغة مارك MARC عن طريق الخاصية MARC Join كما تظهر في الشكل (3-12)، حتى يتم استيراد هذه التسجيلات فيما بعد في النظام الآلي دفعة واحدة بدلا من استيراد كل تسجيلة على حدة.



الشكل رقم (3-12) استخدام أداة MARCJoin لبرنامج مارك إيدت MARC Edit في إعداد ملف واحد يضم مجموعة التسجيلات التي تم استخراجها

4- تم استيراد التسجيلات التي تم تحويلها إلى صيغة مارك في اثنين من النظم الآلية التي تدعم شكل الاتصال مارك 21، النظام الأول هو نظام كوها Koha وهو أحد النظم الآلية المتكاملة والمفتوحة المصدر في مجال المكتبات، والنظام الثاني هو نظام المستقبل للمكتبات Future Library System وهو نظام آلي مصري صادر عن مركز تقنية الاتصالات والمعلومات – جامعة المنصورة.

وتم استخدام نظام كوها من خلال أحد مواقع النسخة التجريبية demo المتاحة مجانا على الإنترنت على الموقع https://koha-community.org/demo/، وبالنسبة لنظام المستقبل تم استخدام النسخة التجريبية demo المتاحة على الإنترنت على الموقع https://eulc.edu.eg/demo. ونجحت عملية الاستيراد في كلا النظامين لجميع الملفات (65 ملف) كما يتبين من الشكلين رقم (3-13)، ورقم (3-14).

دور المكتبات الجامعية السعودية في برامج التعليم عن بعد عبد الرشيد عبد العزيز حافظ

```
000     00881 am a22001213u 4500
042     _adc
100 1   _aعبد الرشيد عبد العزيز حافظ
        _eauthor
245 0 0 _aدور المكتبات الجامعية السعودية في برامج التعليم عن بعد
260     _b,
520     _aتتناول المقالة دور المكتبات الجامعية السعودية في برامج التعليم عن بعد في جامعتي الملك عبد العزيز والإمام محمد بن سعود الإسلامية من وجهة نظر أعضاء هيئة التدريس وعميدي شئون المكتبات في الجامعتين
540     _aجميع الحقوق محفوظة © لمعهد الأمير سلمان للدراسات والخدمات الاستشارية
653     _aالمكتبات الجامعية
787 0   _n1 دراسات المعلومات العدد
```

الشكل رقم (3-13) نموذج لأحد تسجيلات المقالات التي تم استيرادها في نظام كوها Koha



الشكل رقم (3-14) نموذج لأحد تسجيلات المقالات التي تم استيرادها في نظام المستقبل

يتضح مما سبق إمكانية استثمار ما وراء البيانات في عينة التجربة من المقالات، وبالتالي يمكن رفض الفرض الذي ينص على أنه "لا يمكن استثمار ما وراء البيانات في مقالات الدوريات الإلكترونية العربية في مجال المكتبات والمعلومات في إعداد تسجيلات ببليوجرافية وفقا لشكل الاتصال مارك 21 MARC21".

خلاصة الفصل:

1- لاختبار تأثير ما وراء البيانات على كفاءة الاسترجاع من محركات البحث جوجل وياهو وبينج تم إعداد تجربة باستخدام عينة من مقالات الدوريات الإلكترونية العربية المتخصصة في مجال المكتبات والمعلومات بلغ عددها (72) مقالة تنتمي إلى كل من دورية Cybrarians Journal، والعربية 3000، ودراسات المعلومات، والمجلة العربية للدراسات المعلوماتية.

2- صممت التجربة لتشمل ثلاث مجموعات مجموعة ضابطة – المجموعة (أ) تتكون من عينة مقالات التجربة ولا تحتوي على ما وراء البيانات، ومجموعة تجريبية – المجموعة (ب) تحتوي على نفس العينة من المقالات مضاف إليها ما وراء البيانات وفقا لبؤرة دبلن، ومجموعة تجريبية ثانية – المجموعة (ج) تشتمل على عينة المقالات وبها ما وراء البيانات باستخدام مفردات سكيما وإطار وصف المصدر في الخصائص RDFa.

3- تم إتاحة مجموعات التجربة على الإنترنت من خلال ثلاثة مواقع في نفس الوقت هي: www.lisaja.com للمجموعة الضابطة (أ)، وwww.ajainlis.com للمجموعة التجريبية (ب)، وwww.arabiclisja.com للمجموعة التجريبية (ج).

4- تم البحث عن كل مقالة من المقالات عينة التجربة في كل من محرك البحث جوجل، وياهو، وبينج بعد شهرين من نشر مجموعات التجربة على الإنترنت، وذلك باستخدام العنوان، والكلمات المفتاحية.

5- تم استخدام ترتيب ظهور مقالات كل مجموعة في قائمة نتائج محرك البحث كمقياس لكفاءة الاسترجاع، فكلما ظهرت المقالة أولا في قائمة نتائج البحث كلما دل ذلك على تحقيقها لأداء أفضل في الاسترجاع. وتم البحث عن كل مقالة من المقالات عينة التجربة (72 مقالة) بالعنوان والكلمات المفتاحية في محركات البحث محل الدراسة، وعند العثور على المقالة تم تحديد المجموعات وإعطاء كل مجموعة درجة على مقياس من 3 إلى 1، 3 للمجموعة التي تظهر أولا، و2 للمجموعة التي تظهر ثانيا، و1

للمجموعة التي تظهر أخيرا في المرتبة الثالثة، ليتم بعد ذلك تجميع الدرجات التي حققتها مقالات كل مجموعة وإجراء التحليلات والاختبارات الإحصائية.

6- حقق محرك البحث جوجل أكبر معدل للاستدعاء من بين محركات البحث الثلاثة محل الدراسة عند البحث بالعنوان، بما يدل على قدرته العالية وتفوقه في تكشيف ما يتاح على الإنترنت وبخاصة تكشيف النص الكامل. وجاء في المرتبة الثانية محرك البحث بينج، بينما جاء في المرتبة الأخيرة محرك البحث ياهو حيث حقق أقل معدل لاستدعاء المقالات عينة الدراسة.

7- تم قبول فرض الدراسة الذي ينص على أن ما وراء البيانات لا تؤثر على كفاءة استرجاع مقالات الدوريات الإلكترونية العربية في مجال المكتبات والمعلومات في كل من محرك البحث جوجل، وياهو، وبينج سواء عند البحث بالعنوان، أو بالكلمات المفتاحية، مع الأخذ في الاعتبار أن هذه الدراسة ركزت على تأثير ما وراء البيانات المكودة وفقا لبؤرة دبلن، ومفردات سكيما بإطار وصف المصدر RDFa.

8- لوحظ استخدام كل من محرك البحث ياهو وبينج لما وراء البيانات وفقا لبؤرة دبلن، وتحديدا عنصر الوصف Description في توفير ملخص للمقالات عند عرضها في قائمة نتائج محرك البحث.

9- يمكن استثمار ما وراء البيانات الواردة في صفحات العنكبوتية بصفة عامة، والواردة في عينة التجربة من المقالات في إعداد تسجيلات ببليوجرافية وفقا لشكل الاتصال مارك 21 والتي يمكن استثمارها فيما بعد في أحد قواعد البيانات، أو في المستودعات الرقمية، أو في إعداد كشاف لهذه المقالات.

الخاتمــة

النتائج والتوصيات

تمهيد:

سعت هذه الدراسة إلى التعرف على واقع ما وراء البيانات في مقالات الدوريات الإلكترونية العربية في مجال المكتبات والمعلومات من خلال تحليل شفرة المصدر Source Code لعينة من هذه المقالات، كما حاولت اختبار تأثير ما وراء البيانات في صفحات العنكبوتية بشكل عام، وفي مقالات الدوريات الإلكترونية العربية في مجال المكتبات والمعلومات بشكل خاص على كفاءة الاسترجاع من ثلاثة من أبرز محركات البحث وهي: جوجل Google، وياهو Yahoo، وبينج Bing. كما عرّفت بإمكانية استثمار ما وراء البيانات الواردة في هذه المقالات في إعداد تسجيلات ببليوجرافية وفقا لشكل مارك 21. وفيما يلي عرضا لما توصلت إليه الدراسة من نتائج تشكل إجابات على تساؤلاتها، تليها التوصيات التي تقترحها هذه الدراسة.

1 النتائج:

فيما يلي عرضا لنتائج الدراسة كإجابات على التساؤلات التي وضعتها هذه الدراسة:

1 ما الوضع الحالي لما وراء البيانات في مقالات الدوريات الإلكترونية العربية في مجال المكتبات والمعلومات؟

1-1 تتوافر عناصر ما وراء البيانات في جميع مقالات الدوريات الإلكترونية العربية في مجال المكتبات والمعلومات بنسبة 100% من إجمالي المقالات.

2-1 كانت أكثر خطط ما وراء البيانات استخداما في مقالات الدوريات الإلكترونية العربية في مجال المكتبات والمعلومات هي تيجان الميتا HTML Meta Tags، بينما استخدمت بؤرة دبلن في نسبة صغيرة بلغت نحو 2.7% من إجمالي المقالات، وتجدر الإشارة إلى أن جميع هذه المقالات انتمت إلى دورية واحدة فقط هي دورية مجلة المكتبات والمعلومات الصادرة في الجزائر.

3-1 بلغ أقصى عدد من تيجان ما وراء البيانات في مقالات الدوريات الإلكترونية العربية في المجال (11) تاجا، وردت في مقالات مجلة المكتبات والمعلومات، بينما بلغ أقل عدد أربعة (4) تيجان جاءت في جميع مقالات العربية 3000. وكان عدد تيجان ما وراء البيانات في مقالات كل من Cybrarians Journal، ودراسات المعلومات، والمجلة العربية للدراسات المعلوماتية سبعة (7) تيجان.

4-1 بلغ إجمالي أنواع تيجان ما وراء البيانات المستخدمة في إجمالي مقالات الدوريات الإلكترونية العربية في المجال (13) تاجا هي: العنوان title، والوصف description، والكلمات المفتاحية keywords /أو الموضوع subject، والمؤلف author/ أو المنشئ creator، والتاريخ date، واللغة language، والمحدد identifier، والمصدر source، ونوع المحتوى content-type، والمولد generator، والروبوت robot، والعرض viewport، ومدى التوافق X-UA-compatible. ووزعت هذه الأنواع على فئتين أساسيتين هما: عناصر ما وراء البيانات الوصفية، وعناصر ما وراء البيانات الفنية.

5-1 كان تاج العنـوان، والوصـف، والكلـمات المفتاحيـة، ونـوع المحتـوى أكـثر التيجـان استخداما في مقالات الدوريات الإلكترونية العربية في مجال المكتبات والمعلومـات، حيث ورد كل منها في جميع المقالات عينة الدراسة.

6-1 لم تحظى تيجان ما وراء البيانات في مقالات الدوريات الإلكترونية العربية في مجال المكتبات والمعلومات بقدر كاف مـن الجـودة، حيـث وردت أغلـب قـيم التيجـان بشكل غير صحيح لا يتناسب مع وصف كل مقالة من المقالات.

7-1 تعـد مجلـة المكتبـات والمعلومـات التـي تصـدر عـن مخـبر البحـث "تكنولوجيـا المعلومات ودورها في التنمية الوطنية"- الجزائر أكثر الدوريات الإلكترونية العربية المتخصصة في المجال اهتماما بعناصر ما وراء البيانـات وجودتهـا سـواء مـن حيـث عدد التيجان أو العناصر المسجلة لكل مقالة، أو الخطـط المسـتخدمة في ترميزهـا، أو حتى جودة القيم الواردة بها.

وعن وضع ما وراء البيانات في عينة من مقالات الدوريات الإلكترونية الأجنبية في مجـال المكتبات والمعلومات ممثلـة في كـل مـن دوريـة: "D - Lib Magazine"، و"Ariadne"، و"First Monday"، تم التوصل إلى:

توافرت عناصر ما وراء البيانات في جميع مقالات هذه الدوريات، وبلـغ أقصى عـدد مـن تيجان ما وراء البيانات المستخدمة في مقالاتها (18) تاجا وذلك في مقالات دوريـة First Monday.

- مثل الدوريات الإلكترونية العربية في المجال، فقد استخدمت الدوريات الأجنبية في ترميز تيجان ما وراء البيانات في مقالاتها تيجان الميتا HTML Meta Tags، وبـؤرة دبلن. إلا أنها انفـردت باسـتخدام تيجـان هـاي ويربـريس Highwire Press tags التي يوصي بها جوجـل الباحـث العلمـي Google Scholar؛ حيـث اسـتخدمت في حوالي 66.67% من المقالات محل الدراسة وقد انتمت هـذه المقـالات إلى دوريتـين هما Ariadne، وFirst Monday.
- بلغ إجمالي أنواع تيجان أو عناصر ما وراء البيانات الواردة في مقالات الدوريات الإلكترونية الأجنبية (20) عنصرا تشـمل العنـوان، والوصـف، والكلـمات

المفتاحية / الموضوع، والمؤلف، والتاريخ، والمصدر، والعدد issue، والمجلد volume، والترقيم الدولي الموحد للدورية ISSN، واللغة، والمحدد، والمعرف الرقمي للكيان DOI، وموقع ملخص المقالة citation_abstract_html_url، ونوع المصدر type، والحقوق rights، والتغطية coverage، والمساهم contributor، ونوع المحتوى / الشكل، والمولد، ورقم إصدارة عناصر ما وراء البيانات لجوجل الباحث العلمي gs_meta_revision.

- اشتركت كل من مقالات الدوريات العربية والأجنبية في مجال المكتبات والمعلومات في عشرة (10) تيجان شملت: العنوان، والوصف، والكلمات المفتاحية / الموضوع، والمؤلف / المنشئ، والتاريخ، واللغة، والمحدد، والمصدر، ونوع المحتوى، والمولد. بينما انفردت مقالات الدوريات الإلكترونية الأجنبية في فئة عناصر ما وراء البيانات الوصفية بتسعة عناصر لم ترد في مقالات الدوريات الإلكترونية العربية هي: العدد، والمجلد، والترقيم الدولي الموحد للدورية، والمعرف الرقمي للكيان، وموقع ملخص المقالة، ونوع المصدر، والحقوق، والتغطية، والمساهم.
- يعد العنوان، والوصف، ونوع المحتوى أكثر العناصر ورودا في مقالات الدوريات الإلكترونية الأجنبية بنسبة 100% من إجمالي المقالات.
- حرصت مقالات الدوريات الإلكترونية الأجنبية في مجال المكتبات والمعلومات على الجودة في ما وراء البيانات الواردة بها.

2 ما مدى كفاية عناصر بؤرة دبلن في وصف مقالات الدوريات الإلكترونية العربية في مجال المكتبات والمعلومات؟

1-2 توفر مفردات بؤرة دبلن DCMI Terms عناصر ما وراء بيانات يمكن من خلالها وصف مقالات الدوريات الإلكترونية العربية المتخصصة في المجال بشكل كاف، وقد أمكن تحديد هذه العناصر بعشرة عناصر هي:

العنوان Title، والمنشئ Creator، والموضوع Subject، والوصف Description،

والناشر Publisher، والمساهم (إن وجد) Contributor، والتاريخ Date، والمحـدد Identifier، وجزء من IsPartOf، والحقوق Rights.

3 ما مدى كفاية مفردات سـكيما Schema.org في وصـف مقـالات الـدوريات الإلكترونيـة العربية في المجال باستخدام إطار وصف المصدر في الخصائص RDFa؟

1-3 توفر مفردات سكيما Schema.org عنـاصر مـا وراء بيانـات يمكـن اسـتخدامها في وصف مقالات الدوريات الإلكترونية العربية في مجال المكتبات والمعلومات وصـفا مناسبا، وقد تم تحديد عشرة عناصر وهي:

العنـوان الـرئيسي headline، والمؤلـف author، والكلـمات المفتاحيـة keywords، والوصف description، والناشر publisher، والمسـاهم (إن وجـد) contributor، وتاريخ النشر datePublished، والموقع الموحد للمصدر url، وجزء مـن isPartOf، ومالك الحقوق copyrightHolder.

4 هل يؤثر توافر ما وراء البيانات في مقالات الدوريات الإلكترونيـة العربيـة المتخصصـة في المجال على كفاءة استرجاعها من محركات البحث؟

1-4 لاختبار تأثير ما وراء البيانات (العامـل التجريبـي) عـلى كفـاءة اسـترجاع (العامـل التابع) مقالات الدوريات الإلكترونية العربية في مجـال المكتبـات والمعلومـات مـن كل من محرك البحث جوجـل Google، ويـاهو Yahoo، وبيـنج Bing تـم إعـداد تجربة على عينة مـن مقـالات الـدوريات الإلكترونيـة العربيـة في مجـال المكتبـات والمعلومـات بلغـت (72) مقالـة، وانتمـت هـذه المقـالات لأربـع دوريـات هـي: Cybrarians Journal، والعربيـة 3000، ودراسـات المعلومـات، والمجلـة العربيـة للدراسات المعلوماتية.

2-4 اعتمدت التجربة على إنشاء ثلاث مجموعات لعينة التجربة من المقالات وفقـا لمـا يلي: المجموعة الضابطة – المجموعة (أ) تشتمل على صـفحات بلغـة ترميـز الـنص الفائق HTML لعينة التجربة من المقالات بدون إدراج ما وراء البيانات بها، وتـم إتاحتها على الموقع www.lisaja.com .

المجموعة التجريبية الأولى – المجموعة (ب) تشتمل على نفس الصفحات مضافا إليها ما وراء البيانات وفقا لبؤرة دبلن. وتم إتاحتها على الموقع www.ajainlis.com .

المجموعة التجريبية الثانية – المجموعة (ج) تشتمل على نفس الصفحات مضافا إليها ما وراء البيانات وفقا لمفردات سكيما Schema.org باستخدام إطار وصف المصدر في الخصائص RDFa. وأتيحت على الموقع www.arabiclisja.com .

3-4 تم البحث عن كل مقالة من المقالات عينة التجربة في كل محرك من محركات البحث باستخدام العنوان، والكلمات المفتاحية. واستخدم ترتيب الظهور الذي تحققه مقالات كل مجموعة في قائمة نتائج البحث لمحركات البحث كمقياس لكفاءة الاسترجاع الذي تحققه كل مجموعة. وعلى مقياس من ثلاثة إلى واحد تم إعطاء درجات لمقالات كل مجموعة، فالمجموعة التي تظهر إحدى مقالاتها أولا في قائمة نتائج البحث تأخذ الدرجة (3)، في حين أن المجموعة التي تظهر مقالاتها في المرتبة الثانية تأخذ الدرجة (2)، والتي تظهر أخيرا تأخذ الدرجة (1).

4-4 لا تؤثر ما وراء البيانات على كفاءة استرجاع مقالات الدوريات الإلكترونية العربية من كل من محرك البحث جوجل، وياهو، وبينج وذلك وفقا لنتائج الاختبار الإحصائي كروسكال ويلز الذي تم استخدامه في قياس فروض الدراسة.

5-4 حقق محرك البحث جوجل أكبر معدل للاستدعاء من بين محركات البحث الثلاثة عند البحث بالعنوان، بما يدل على قدرته العالية وتفوقه في تكشيف ما يتاح على الإنترنت وبخاصة تكشيف النص الكامل، تلاه محرك البحث بينج، وجاء في المرتبة الأخيرة محرك البحث ياهو الذي حقق أقل معدل لاستدعاء المقالات عينة الدراسة.

6-4 تفوق محرك البحث بينج على غيره من محركات البحث عينة الدراسة في معدل الاستدعاء عند البحث بالكلمات المفتاحية ؛ حيث بلغ أقصى عدد للمقالات

التي استرجعها لمجموعة من مجموعات التجربـة (21) مقالـة تنتمـي للمجموعـة التجريبية (ج) وبنسبة حوالي 29.16% تلاه محرك البحث ياهو الذي كانـت أعـلى نسبة استدعاء يحققها هي 25 % لمقالات المجموعة التجريبيـة (ب)، بيـنما حقـق جوجل معدل استدعاء أقل بلغ 11.11% للمجموعة الضابطة (أ). وهـذا يشـير إلى تفوق محرك البحث بيـنج في التكشـيف وفقـا للمعنـى ودلالـة استفسـار البحـث وليس فقط تكشيف النص الكامل.

7-4 على الرغم من عدم تأثير ما وراء البيانات على كفـاءة الاسـترجاع لكـل مـن محـرك البحث ياهو، وبينج إلا أن كلا من المحركين اسـتخداما مـا وراء البيانـات وتحديـدا عنصر الوصف Description المكود وفقا لبؤرة دبلن في تـوفير ملخـص للمقـالات عند عرضها في قائمـة النتـائج. كـما أن كـلا مـن المجمـوعتين التجـريبيتين (ب، وج) حققتا معدل اسـتدعاء كـان دائمـا أعـلى مـن المجموعـة الضـابطة (أ) وذلـك عنـد البحث في كل من محرك البحث ياهو، وبينج. وفي هذا إشارة إلى أهمية نسبية لمـا وراء البيانات في هذين المحركين.

8-4 على الرغم من أن نتائج الدراسة تشير إلى عدم تأثير ما وراء البيانات على محركات البحث، إلا أن ذلك ينطبق بصفة محددة على ما وراء البيانـات التـي تـم ترميزهـا باستخدام خطط بعينها هي بؤرة دبلن، ومفردات سكيما.

5 هل يمكن استثمار ما وراء البيانات في مقالات الـدوريات الإلكترونيـة العربيـة في المجـال واستخلاصها؟

1-5 من خـلال إعـداد برنـامج لاسـتخلاص مـا وراء البيانـات مـن المقـالات عينـة الدراسة، أمكن إعداد تسجيلات ببليوجرافية وفقا لشـكل الاتصـال مـارك 21 MARC 21 وتـم اسـتيرادها وإيـداعها بنجـاح في اثنـين مـن الـنظم الآليـة المتخصصة في المكتبات، وهي نظام كوها Koha، ونظـام المسـتقبل Future Library System. وتؤكد هذه النتيجة على إمكانية استثمار ما وراء بيانات مقالات الدوريات الإلكترونية العربية في مجال المكتبات والمعلومـات بصـفة

خاصة، وكذلك إمكانية استثمار ما وراء البيانات في صفحات العنكبوتية بصفة عامة.

وتنفي هذه النتيجة فرض الدراسة: "لا يمكن استثمار ما وراء البيانات في مقالات الدوريات الإلكترونية العربية لإعداد تسجيلات فهرسة وفقا لشكل مارك MARC 21".

2 التوصيات:

في ضوء ما توصلت إليه الدراسة من نتائج، يمكن الإشارة إلى التوصيات التالية:

1 عند تسجيل ما وراء البيانات في مقالات الدوريات الإلكترونية العربية في مجال المكتبات والمعلومات يقترح اتباع ما يلي:

1-1 يمكن الاكتفاء بإدراج عناصر ما وراء البيانات الأساسية لكل مقالة والتي تتمثل في: العنوان، والوصف، والكلمات المفتاحية، خاصة إذا كان إدراج مزيد من العناصر لكل مقالة يشكل عبئا إضافيا على الدورية من حيث الوقت، والجهد. وذلك وفقا لما يلي:

<title>عنوان المقالة</title>

<meta name="description" content="وصف أو ملخص المقالة ">

<meta name="keywords" content="الكلمة المفتاحية الأولى ,الكلمة المفتاحية الثانية">

2-1 ضرورة إدراج عناصر ما وراء البيانات بشكل صحيح يتناسب مع كل مقالة على حدة، وبشكل خاص كل من عنصر العنوان، والوصف حيث يتم استخدام كل منهما عند عرض هذه المقالات في قائمة نتائج محركات البحث.

3-1 يمكن الاستعانة بالبرامج الآلية، سواء المتاحة مجانا على الإنترنت، أو تلك المتاحة كإضافات Plugins لبرامج إدارة محتوى المواقع CMS في إدراج ما وراء البيانات بشكل صحيح. ومن هذه البرامج ما يلي:

- أداة Free Meta Tag Tool المتاحة على موقع SEO Book (http://tools.seobook.com/meta-medic/).

- أداة Meta Tags Generator المتاحـــة عـــلى موقـــع Web Code (/http://webcodetools.com/meta-tags-generator)Tools.
- أداة Online Meta Tag Generator المتاحة على موقع OnlineMetaTag /(http://www.onlinemetatag.com)

4-1 يمكن لكل دوريـة أن تخصـص عـددا مـن الأفـراد لتـولي مسـؤولية إدراج مـا وراء البيانات بشكل صحيح لكل مقالة من مقـالات الدوريـة وذلـك بخـلاف المشرفـون على تصميم ونشر المقالات وإتاحتها على الإنترنت.

5-1 يمكن الاستعانة بأداة جوجل Data Highlighter لتحسين عرض المقالات في قائمـة نتائج بحث جوجل.

6-1 الاكتفاء بإدراج عناصر ما وراء البيانات وفقا لتيجـان الميتـا HTML Meta Tags، حيث ثبت من تجربة الدراسة أن إدراج عناصر مـا وراء البيانـات باسـتخدام بـؤرة دبلن أو مفردات الاسكيما باستخدام RDFa لا يؤثر على محركات البحث.

7-1 يمكن للدوريات الإلكترونية إدراج عناصر ما وراء البيانات باستخدام التيجان التـي أوصى بهـا محـرك البحـث جوجـل العلمـي Google Scholar، والتـي تمكـن هـذا المحرك بصفة خاصة من تكشيف وعـرض هـذه المقـالات بالشـكل الـذي يتناسـب معه. ويمكن الاستعانة بالمثال التالي:

<meta name="citation_title" content="عنوان المقالة">

<meta name="citation_author" content="المؤلف الأول">

<meta name="citation_author" content="المؤلـف الثـاني إن وجـد، ويكـرر هـذا العنصر إذا كان هناك مؤلفين آخرين">

<meta name="citation_publication_date" content="تاريخ النشر بـاليوم والشـهر والسنة">

<meta name="citation_journal_title" content="عنوان الدورية">

<meta name="citation_volume" content="رقم المجلد">

<meta name="citation_issue" content="رقم العدد">

<meta name="citation_firstpage" content="رقم الصفحة الأولى من المقالة">

<meta name="citation_lastpage" content="رقم الصفحة الأخيرة من المقالة">

<meta name="citation_pdf_url" content="الموقع الموحد للمقالة بصيغة pdf">

2 على أقسام المكتبات والمعلومات الانتباه إلى أن ما وراء البيانات كانت في وقت سابق تؤثر على كفاءة الاسترجاع من محركات البحث ولكن في الوقت الراهن وفقا لما توصلت له هذه الدراسة فإنه لا يوجد تأثير لها على كفاءة الاسترجاع وبخاصة تلك المرمزة وفقا لبؤرة دبلن وضرورة مراعاة ذلك في المقررات الدراسية المتعلقة بما وراء البيانات.

3 الحاجة إلى زيادة عدد الدوريات الإلكترونية العربية المتخصصة في المجال، من خلال تشجيع الناشرين والمؤسسات الأكاديمية على ذلك.

4 الحاجة إلى إنشاء أرشيف للدوريات الإلكترونية العربية في مجال المكتبات والمعلومات على غرار أرشيف الإنترنت نظرا للتغير السريع في مواقع هذه الدوريات.

5 يمكن أن تتيح الدوريات الإلكترونية المتخصصة في المجال أمام المستفيدين منها إمكانية التوسيم Tagging للمقالات التي تتاح عليها.

6 في إطار ما توصلت له الدراسة من إمكانية استثمار ما وراء البيانات الواردة في صفحات العنكبوتية بصفة عامة والواردة في مقالات الدوريات الإلكترونية العربية في المجال بصفة خاصة، يمكن لأحد المكتبات الكبرى مثل مكتبة الإسكندرية، أو الفهرس العربي الموحد إنشاء محرك بحث عربي يتاح من خلال مواقع هذه المكتبات ويعتمد في خوارزميته على ما وراء البيانات، على أن يهتم هذا المحرك بالدرجة الأولى بمصادر المعلومات العلمية – على غرار جوجل الباحث العلمي Google Scholar – وذلك في مقابل محركات البحث المتاحة على الإنترنت مثل جوجل وبينج التي تُعنى بكل ما يتاح على الإنترنت.

الدراسات المستقبلية:

1. دراسة المحتوى العربي المتاح من خلال جوجل الباحث العلمي Google Scholar.
2. اختبار تأثير ما وراء البيانات وفقا لتيجان الميتا HTML Meta Tags على أداء محركات البحث.
3. اختبار تأثير ما وراء البيانات وفقا لطرق ترميز أخرى مثل المايكروداتا Microdata على أداء محركات البحث.

قائمة المصادر

قائمـة المصادر

أولا: المصادر العربية

أبـو النـور، إينـاس محمـد فـوزي. تقنيـات البيانـات الخلفيـة لوصـف المصـادر الإلكترونيـة واسترجاعها على شبكة الإنترنت: دراسة ميدانية مقارنة / إشراف مصطفي حسام الدين. أطروحة (ماجستير). جامعة القاهرة، كلية الآداب، قسم المكتبات والوثائق والمعلومات، 2008.

بامفلح، فاتن. الميتاداتا وتنظيم مصادر المعلومات الإلكترونية في المكتبات. دراسـات عربيـة في المكتبات وعلم المعلومات. مج 7، ع 3. (2002). ص ص 24 – 54.

البسيوني، بدوية محمد. تيجان الميتاداتا meta tags ومدى تمثيلها في صفحات الويب: دراسـة تطبيقية عـلى مواقـع التميـز الرقمـي العربيـة عـلى الانترنـت. الاتجاهـات الحديثـة في المكتبات والمعلومات. مج 16، ع32 (يوليو 2009). ص ص 179 – 224.

تيلور، ألين ج. تنظيم المعلومات. جزء 1. / تأليف ألين ج. تيلـور، دانيـل ن. جـودري ؛ ترجمـة هاشم فرحات. الرياض: جامعة الملك سعود، النشر العملي والمطابع،2012. ص374.

الجندي،،محمود عبد الكريم. الاتجاهات الحديثة في الميتاداتا: مراجعة علمية للإنتاج الفكري. الاتجاهات الحديثة في مجال المكتبات والمعلومـات. مـج 14، ع 27. (ينـاير 2007). ص ص 81- 130.

جوهري، عزة فاروق. الميتاداتا ودعم استرجاع المحتوى الرقمي للصحف العربية الإلكترونية في البيئة الرقمية: دراسة تطبيقية لمدى تمثيلها في بعض الصحف المصرية والسعودية. بحوث في علم المكتبات والمعلومات. ع 7 (سبتمبر 2011). ص ص 45 – 91.

خليفة، محمود عبد الستار. مواقع الإنترنت العربية في مجال المكتبات والمعلومات: دراسة تحليلية [على الخط المباشر]. Cybrarians Journal. ع 4 (مارس 2005). متاح على: http://www.journal.cybrarians.info/index.php?option=com_content&view=article&id=505:2011-08-16-21-36-31&catid=241:2011-08-16-08-16-55 . [تاريخ الزيارة 2015/4/25].

خيرو، مازن عمر. مواقع تحليل بيانات الميتاداتا على شبكة الإنترنت: دراسة تحليلية . دراسات المعلومات. ع 3 (سبتمبر 2008) .ص ص 61 – 77.

الزهيري، طلال ناظم. البيانات الفوقية للمواقع الحكومية العراقية على الإنترنت وتأثيرها في آلية تكشيفها في محركات البحث [على الخط المباشر]. المجلة العراقية لتكنولوجيا المعلومات. مج 1، ع 2 (2008). متاح على: http://drtazzuhairi.blogspot.com.eg/2011/12/blog-post.html . [تاريخ الزيارة 2014/1/21].

شاهين، شريف كامل. واصفات البيانات (Metadata) مصدرا لتسجيلات الفهرسة القياسية لمصادر المعلومات الإلكترونية الشبكية العربية: دراسة استكشافية تجريبية. الاتجاهات الحديثة في المكتبات والمعلومات. مج 9، ع 18 (يوليو 2002).ص ص 79- 146.

عبد الجواد، سامح زينهم. خطط ومعايير ميتاداتا الحفظ الرقمي: دراسة تحليلية.1. مجلة المكتبات والمعلومات العربية. س 30، ع 1(يناير 2010). ص ص 163 - 196.

عبد الفتاح، خالد. The Effect of Metadata in Discovering Web Resources. في: ندوة فهرسة مصادر الإنترنت واستخدام معايير الميتاداتا ودبلن كور (4-8 يوليو 2004). 21 ص. القاهرة: المنظمة العربية للتنمية الإدارية، جامعة الدول العربية.

عبد الهادي، زين. وصائف البيانات (Metadata) في مواقع المكتبات العربية في كل من مصر

والسعودية [عـلى الخـط المبـاشر]: دراسـة تطبيقيـة. . Cybrarians Journal ع 4 (مـارس 2005). متـاح عـلى: http://www.journal.cybrarians.info/index.php?option=com_content&view=article&id=503:metadata&catid=241:2011-08-16-08-16-55 . [تاريخ الزيارة 2015/4/2].

عبد الهادي، محمد فتحي. الميتاداتا: أسسها النظرية وتطبيقاتها العملية / محمد فتحـي عبـد الهادي، خالد عبد الفتاح. القاهرة: الدار المصرية اللبنانية، 2013. 327ص.

العربي، أحمد عبادة. الميتاداتا ودورها في دعم المحتوى الرقمي: دراسـة تطبيقيـة عـلى مواقـع دوريات الوصول الحر في المكتبات والمعلومات على الإنترنت. مجلة مكتبـة الملـك فهـد الوطنية. مج 16، ع 1. ص ص 131 – 132.

عكوش، نبيل، غانم، نذير، كوكة، سليمة . ما وراء البيانات ومعاييرها في بيئة المكتبـات [عـلى الخـط المباشر]. في: المؤتمر الدولي لتكنولوجيا المعلومات الرقمية: الاتجاهات الحديثـة في تكنولوجيـا المعلومـــات. (عـــمان:جامعـــة الزرقـــاء: 9 – 11 أكتـــوبر 2012) . متـــاح عـــلى: http://digital.jilwan.com/digital2012/download2012.php?f=jalsa1/1_4.pdf . [تاريخ الزيارة 2016/1/30].

فلمبان، سوزان بنت مصطفى بن عباس. مدونات المكتبات الجامعية: دراسة تحليلية في دعـم برمجيات المدونات للميتاداتا (واصـفات البيانـات) / إشراف محمـد أمـين عبـد الصـمد مرغلاني، بدوية محمد البسيوني. أطروحة (دكتوراه). جامعة الملـك عبـد العزيـز، كليـة الآداب والعلوم الإنسانية، قسم المعلومات، 2011

قاســم، حشــمت. الــدوريات الإلكترونيــة التخصصــية: تطورهــا و تحــدياتها الاجتماعيــة والاقتصــادية. ص ص 333 – 367. في: حشــمت قاســم . الاتصــال العلمــي في البيئــة الإلكترونية . القاهرة: دار غريب للطباعة و النشر و التوزيع، 2005.

كابلن، برسيلا. أساسيات ما وراء البيانات لاختصاصيي المكتبات والمعلومات / ترجمة وتعليـق هاشم فرحات. الرياض: مكتبة الملك فهد الوطنية، 2007. مطبوعات مكتبة الملـك فهـد الوطنية. السلسلة الأولى، ص 54.

لانكستر، ولفرد. نظم استرجاع المعلومات / ترجمة حشمت قاسم. القاهرة: دار غريب، 1981. ص527.

المرغلاني، محمد أمين عبد الصمد، سوزان فلمبان. الميتاداتا في المواقع الحكومية: دراسة تحليلية. دراسات المعلومات، ع 2 (مايو 2008).

النشرتي، مؤمن سيد. محركات البحث الدلالية على الشبكة العنكبوتية: دراسة مسحية تحليلية لوضع مواصفات محركات بحث دلالية عربية / إشراف أسامة السيد محمود. أطروحة (ماجستير) - جامعة القاهرة، كلية الآداب، قسم المكتبات والوثائق وتقنية المعلومات، 2012.

هندي، عبد الله هندي. دور الميتاداتا في استرجاع مصادر المعلومات الإلكترونية العربية المتاحة على شبكة الإنترنت: دراسة وصفية تحليلية/ إشراف زين الدين محمد عبدالهادي، نوال محمد عبدالله. أطروحة (ماجستير) - جامعة حلوان، كلية الآداب، قسم المكتبات والمعلومات، 2012.

الهيئة المصرية العامة للمواصفات والجودة. 5146 / 2014: معلومات وتوثيق - مجموعة عناصر بيانات دبلن كور للميتاداتا (البيانات الوصفية)، 2014 [على الخط المباشر]. متاح على: http://www.eos.org.eg/ar/standard/4125 . [تاريخ الزيارة 2016/5/27].

ثانيا: المصادر الأجنبية

Aghaei, Sareh, Nematbakhsh, Mohammad Ali and Farsani, Hadi Khosravi. Evolution of the World Wide Web from Web 1.0 to Web 4.0 [online]. International Journal of Web & Semantic Technology (IJWesT). Vol.3, No.1(January 2012). Available at: www.airccse.org/journal/ijwest/papers/3112ijwest01.pdf. [Accessed 8/2/2016]

Alexa Top 500 Global Sites [online] / Alexa. Available at: http://www.alexa.com/topsites. [Accessed 28/2/2016].

Alimohammadi , D. Measurement of the Presence of keywords and description meta-tags on a selected number of Iranian Websites . Online Information Review. Vol. 28, No.3 (2004). Pp..220-223.

Berners-Lee, T, Connolly, D. Hypertext Markup Language - 2.0 [online]. November 1995. Available at: http://tools.ietf.org/html/rfc1866. [Accessed 13/2/2016].

Berners-Lee, T.; Hendler, J. & Lassila, O.: The Semantic Web .Scientific American. Vol. 284, No. 5 (May 2001). Pp. .34-43. As cited in: Brase, Jan. O. Usage of metadata [online]. (PhD) - Elektrotechnik und Informatik der Universt at Hannover. 2005.117p. Available at: https://pdfs.semanticscholar.org/ea90/19ad62ce1ab487881610fe393e9d0c4da6ce.pdf . [Accessed 29/11/2013].

Berners-Lee, Tim. Axioms of Web Architecture: Metadata [online]. 1997. Available at: https://www.w3.org/DesignIssues/Metadata.html . [Accessed 1/2/2016].

Brickley, Dan and Miller, Libby. FOAF Vocabulary Specification 0.99 [online]. Available at: http://xmlns.com/foaf/spec/ .[Accessed 29/3/2016].

BlueGriffon [online]. Available at: http://bluegriffon.org/ .[Accessed 8/11/2014]

Calgary & Konia, B.S. Search Engine Optimization with WebPosition GOLD 2. Plano, Texas: Wordware Publishing , 2002.

Choudhury, Nupur. World Wide Web and Its Journey from Web 1.0 to Web 4.0.[online]. (IJCSIT) International Journal of Computer Science and Information Technologies. Vol. 5, No. 6 (2014). Available at: www.ijcsit.com/docs/Volume%205/vol5issue06/ijcsit20140506265.pdf . [Accessed 8/2/2016].

DCMI Metadata Basics [online]. Available at: http://dublincore.org/metadata-basics/ . [Accessed 19/10/2013].

DCMI Metadata Terms [online]. 14-06-2012. Available at: http://dublincore.org/documents/2012/06/14/dcmi-terms/ . [Accessed 20/9/2014].

DCMI Specifications [online]. Available at: http://dublincore.org/specifications/ . [Accessed 19/10/2013]

Dublin Core [online]. In: Wikipedia: The Free Encyclopedia. Available at: https://en.wikipedia.org/wiki/Dublin_Core .[Accessed 20/6/2014].

Dublin Core Metadata Element Set, Version 1.1[online]. 14-06-2012. Available at: http://www.dublincore.org/documents/dces/ .[Accessed 21/3/2016].

Dublin Core Qualifiers [online]. 11-07-2000. Available at: http://dublincore.org/documents/2000/07/11/dcmes-qualifiers/ .[Accessed 22/10/2013].

Ellett, Robert O. Internet Search Engines giving you Garbage [online]. Available at: http://scis.nova.edu/ellettro .[Accessed 19/10/2013].

Elser, J.K. & Paxton, J..A Search Engine That Learns. In: Proceedings of the Fourth International Association of Science and Technology for Development (IASTED) International Conference on Computational Intelligence, July 2005 Canada.

Existing rel values. Microformats Wiki. [online] Available at: http://microformats.org/wiki/existing-rel-values#HTML5_link_type_extensions. [Accessed 24/6/2016].

Expressing Dublin Core metadata using HTML/XHTML meta and link elements [online]. Available at: http://dublincore.org/documents/dc-html/. [Accessed 18/8/2015].

Facebook. Open Graph Markup [online]. Available at: https://developers.facebook.com/docs/sharing/webmasters#markup .[Accessed 23/3/2016].

Gill, Tony. Metadata and the Web [online]. In: Introduction to Metadata. Online Edition, Version 3.0. Los Angeles: Getty Research Institute, 2008. Available at: http://www.getty.edu/research/publications/electronic_publications/intrometadata/metadata.html. [Accessed 25/11/2013].

Gilliland, Anne J . Setting the Stage [online]. In: Introduction to Metadata. Online Edition, Version 3.0. Los Angeles: Getty Research Institute, 2008. Available at: http://www.getty.edu/research/publications/electronic_publications/intrometadata/setting.html . [Accessed 25/11/2013].

Giomelakis, Dimitrios,Veglis, Andreas. Employing Search Engine Optimization Techniques in Online News Aritcles [online]. Studies in Media and Communication. Vol. 3, No. 1 (June 2015). Available at:: http://dx.doi.org/10.11114/smc.v3i1.683 . [Accessed 26/11/2015].

Google Developers. Providing Structured Data. [online] Available at: https://developers.google.com/custom-search/docs/structured_data#using-meta-tags.[Accessed 22/10/2015].

Google Developers. Structured Data Testing Tool [online]. Available at: https://developers.google.com/structured-data/testing-tool/ . [Accessed 7/4/2016].

Google Search [online]. In: Wikipedia: The Free Encyclopedia. Available at: https://en.wikipedia.org/wiki/Google_Search .[Accessed 20/12/2015].

Google. About Data Highlighter - Search Console Help [online]. Available at: https://support.google.com/webmasters/answer/2692911?hl=en .[Accessed 10/4/2016].

Greenberg, Jane. Metadata and the World Wide Web [online]. In: Encyclopedia of Library and Information Science. Vol. 3.2nd ed. p1876. Available at: https://books.google.com.eg/books?id=Sqr-_3FBYiYC&printsec=front cover&hl=ar#v=onepage&q&f=false .[Accessed 8/2/2016].

Henshaw, Robin & Valauskas, Edward J. Metadata as a Catalyst Experiments with Metadata and Search Engines in the Internet Journal First Monday [online]. Libri. Vol. 51 (2001). Pp.86-101. Available at: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.102.6248&rep =rep1&type=pdf . [Accessed 22/6/2016].

Henshaw, Robin. The First Monday Metadata Project [online] . Libri, Vol. 49 (1999). Pp.125–131. Available at: http://www.degruyter.com/dg/viewarticle.fullcontentlink:pdfeventlink/$002fj$00 2flibr.1999.49.issue-3$002flibr.1999.49.3.125$002flibr.1999.49.3.125.pdf/libr.1999. 49.3.125.pdf?t:ac=j$002flibr.1999.49.issue-3$002flibr.1999.49.3.125$002flibr.1999 .49.3.125.xml.[Accessed 30/5/2016].

History of the World Wide Web [online]. In: Wikipedia: The Free Encyclopedia. Available At: https://en.wikipedia.org/wiki/History_of_the_World_Wide_ Web .Accessed [8/2/2016].

HTML [online]. In: Wikipedia: the free encyclopedia. Available at: https://en.wikipedia.org/wiki/HTML . [Accessed 13/2/2016].

Jisc Digital Media | Guide [online]: Putting Things in Order: a Directory of Metadata Schemas and Related Standards. Available at: http://www. jiscdigitalmedia.ac.uk/guide/putting-things-in-order-links-to-metadata-schemas-and-related-standards . [Accessed 18/1/2015].

Kinsella, Sheila. Augmenting Social Media Items with Metadata using Related Web Content [online]/ Supervisor John Breslin. Digital Enterprise Research Instiitute, National University of Ireland, Galway, January 2012 (Ph.D.). Available at: https://aran.library.nuigalway.ie/xmlui/bitstream/handle/10379/2674/pap er.pdf?sequence=1 .[Accessed 17/2/2016].

Koehler, Wallace, A Call to Action: What Every Searcher Should Know – And Do About Domain Names, Standards, and metadata. Searcher. Vol. 10,

No 9. (Oct. 2002)..

Kok , Ruud. The Value of Search Engines Ranking [online]: The Effect of Search Engine Ranking on a Web Site's Added Value. (Master Thesis). university of Twente, school of Management and Governance, business administration .2008. 51p. Available at: http://essay.utwente.nl/59288/1/scriptie_R_Kok.pdf . [Accessed 17/3/2013].

Meta element [online]. In: Wikipedia: The Free Encyclopedia .Available at: https://en.wikipedia.org/wiki/Meta_element . [Accessed 11/9/2015].

Meta Tag Analysis Tool by MetaChecker.net [online]. Available at: http://www.metachecker.net/ . [Accessed 15/11/2015].

Meta tags - Custom Search Help [online]. Available at: https://support.google.com/customsearch/answer/2595557?hl=en . [Accessed 27/5/2014]

Metadata Definition .In: TechTerms. Available at: http://techterms.com/definition/metadata . [Accessed 1/2/2016].

Metadata[online].In: Wikipedia: The Free Encyclopedia. Available at: https://en.wikipedia.org/wiki/Metadata . [Accessed 1/9/2015].

Okerson , Ann. The Electronic Journal: what , whence and when ? .The Public Access Computer Systems Review. Vol.2 , No.1 (1991), 5-24.

PageRank [online]. In: Wikipedia: The Free Encyclopedia. Available at: https://en.wikipedia.org/wiki/PageRank . [Accessed 20/12/2015].

Patel, Karan. Incremental Journey for World Wide Web: Introduced with Web 1.0 to Recent Web 5.0 – A Survey Paper [online]. International Journal of Advanced Research in Computer Science and Software Engineering. Vol. 3, No. 10 (October 2013). Available at: http://www.ijarcsse.com/docs/papers/Volume_3/10_October2013/V3I10-0149.pdf . [Accessed 8/2/2016].

Petiya, Sean. Building a Semantic Web of Comics: Publishing Linked Data in HTML/RDFa [online]. (Master Thesis) - Kent State University - College of Communication and Information.2014. 205 p. Available at: http://rave.ohiolink.edu/etdc/view?acc_num=kent1416791055.[Accessed 29/11/2014].

Raosoft. Sample Size Calculator by Raosoft, Inc.[online]. Available at: http://www.raosoft.com/samplesize.html.[Accessed 24/2/2015].

RDFa play [online]. Available at: https://rdfa.info/play/

Reitz, Joan M. ODLIS [online]: Online Dictionary for Library and Information Science. Available at http://www.abc-clio.com/ODLIS/odlis_about.aspx . [Accessed 29/1/2016].

Riley, Jenn. Seeing Standards: A Visualization of the Metadata Universe [online] / content Jenn Riley ; design Devin Becker. 2009 – 2010. Available at: http://www.dlib.indiana.edu/~jenlrile/metadatamap/seeingstandards.pdf . [Accessed 26/12/2015].

Ronallo, Jason. HTML5 Microdata and Schema.org. [online].The Code4lib Journal. No. 16 (03/02/2012). Available at: http://journal.code4lib.org/articles/6400 . [Accessed 29/5/2014].

Safari,Mehdi. Search Engines and Resource Discovery on the Web: Is Dublin Core an Impact Factor?[online]. Webology. Vol. 2, No. 2 (August 2005). Available at: http://www.webology.org/2005/v2n2/a13.html. [Accessed 20/1/2016].

Schema.org . FAQ - schema.org [online]. Available at: http://schema.org/docs/faq.html .[Accessed 18/10/2015].

Schema.org . Schemas - schema.org. [online]: Organization of Schemas. Available at: http://schema.org/docs/schemas.html .[Accessed 18/10/2015].

Schema.org. Getting Started - schema.org [online]. Available at: https://schema.org/docs/gs.html . [Accessed 27/5/2016].

Schema.org. Home - schema.org [online]. Available at: http://schema.org/ [Accessed 18/10/2015].

Schema.org. ScholarlyArticle - schema.org. [online]. Available at: http://schema.org/ScholarlyArticle. [Accessed 18/10/2015].

Smith-Yoshimura, Karen, Shein, Cyndi. Social Metadata for Libraries, Archives and Museums [online]. part 1: Site Reviews. Dublin, Ohio:OCLC Research, 2011. Available at: http://www.oclc.org/research/publications/library/2011/2011-02.pdf . [Accessed 9/2/2016].

Social graph [online]. In: Wikipedia: The Free Encyclopedia. Available at: https://en.wikipedia.org/wiki/Social_graph . [Accessed 16/2/2016].

Social Media Metadata: Facebook Open Graph and Twitter Cards [online]. Available at: https://www.websightdesign.com/services/internet-marketing/search-engine-optimization/social-media-metadata. [Accessed 17/2/2016].

Starr, Barbara. Google's Hunger For Structured Markup [online]. Search Engine Land, 2013. Available at: http://searchengineland.com/googles-hunger-for-structured-markup-160240 . [Accessed 5/4/2016].

Statistic (June 4, 1997), META attributes by count, Vancouver Webpages, Available at: http://vancouver-webpages.com/META/bycount.shtml. [Accessed June 3, 2007]. As cited In: Meta element [online].In: Wikipedia: the free encyclopedia. Available at: https://en.wikipedia.org/wiki/Meta_element . [Accessed 11/9/2015].

Thakur, Dinesh. What is Metadata OR Data Dictionary?. In: eComputer Notes. Available at: http://ecomputernotes.com/fundamental/what-is-a-database/metadata-or-data-dictionary .[Accessed 6/2/2016].

The Advanced Dublin Core Generator! [online]. Available at . http://www.dublincoregenerator.com/generator.html .[Accessed 10/8/2015].

The Open Graph protocol [online]. Available at: http://ogp.me/. [Accessed 7/9/2014].

The Semantic Web in Action [online] / Lee Feigenbaum ... [et. al.]. Scientific American. (19 Jan. 2009). Available at: http://www.scientificamerican.com/article/semantic-web-in-actio/. [Accessed 20/2/2014].

The W3C Markup Validation Service [online]. Available at: https://validator.w3.org/.

Top 10 Search Engines [online]. SEO Consultants Directory . Available at: http://www.seoconsultants.com/search-engines/ [Accessed 22/4/2011].

Twitter. Summary Card [online]. Available at: https://dev.twitter.com/cards/types/summary .[Accessed 24/3/2016].

Twitter. Twitter Cards [online]. Available at: https://dev.twitter.com/cards/overview . [Accessed 24/3/2016].,

User Guide - DCMI_MediaWiki [online].Available at: http://wiki.dublincore.org/index.php/User_Guide. [Accessed 10/10/2013].Using hashtags on Twitter. Help Center. Available at: https://support.twitter.com/articles/49309?lang=en . [Accessed 16/2/2016].

Visser, EB , Kritzinger , WT.& Weideman , M . (2006) . Search Engine Optimizing Elements and Their Effect on Website Visibility: Implementation of the Chambers Model [online] . In: Proceedings Of The 8th Annual Conference On World Wide Web Applications . Cape Town:

Cape Peninsula University of Technology. Available at : https://www.researchgate.net/profile/Melius_Weideman/publication/242405903_Search_engine_optimising_elements_and_their_effect_on_Website_visibility_implementation_of_the_Chambes_model/links/00b49530b29b428a86000000.pdf . [Accessed 22/6/2016].

W3C. Document metadata — HTML5 [online].Available at: http://www.w3.org/TR/html5/document-metadata.html#the-meta-elemen . [Accessed 12/10/2015].

W3C. Good Ontologies - W3C Wiki [online]. Available at: https://www.w3.org/wiki/Good_Ontologies . [Accessed 29/3/2016].

W3C. Linked Data Glossary [online]: W3C Working Group Note 27 June 2013. Available at: https://www.w3.org/TR/2013/NOTE-ld-glossary-20130627/ . [Accessed 30/3/2016].

W3C. Ontologies - W3C [online]. Available at: https://www.w3.org/standards/semanticweb/ontology .[Accessed 27/6/2014].

W3C. RDF Primer [online]:W3C Recommendation 10 February 2004. Available at: https://www.w3.org/TR/2004/REC-rdf-primer-20040210/#basicconcepts . [Accessed 28/12/2013].

W3C. RDFa 1.1 Primer - Third Edition [online]: Rich Structured Data Markup for Web Documents . 2015. Available at: https://www.w3.org/TR/rdfa-primer/ . [Accessed 20/6/2015].

W3C. RDFa Core 1.1 - Third Edition [online]: Syntax and processing rules for embedding RDF through attributes. 2015. Available at: https://www.w3.org/TR/rdfa-core/ . [Accessed 20/6/2015].

W3C. RDFa Lite 1.1 - Second Edition [online]. 2015. Available at: https://www.w3.org/TR/rdfa-lite/ . [Accessed 20/6/2015].

W3C. Semantic Web [online]. Available At: https://www.w3.org/2001/sw/wiki/Main_Page . [Accessed 27/2/2016].

W3C. W3C HTML [online]: The Web's Core Language Available at: https://www.w3.org/html/ . [Accessed 30/3/2016].

W3C.RDFa. [online]: Linked Data in HTML.. Available at: http://rdfa.info/ .[Accessed 2/4/2016].

W3C.What's New in RDF 1.1 [online]. Available at: http://www.w3.org/TR/rdf11-new/ .[Accessed [26/3/2016].

W3schools.com. HTML <meta> Tag [online]. Available at:

http://www.w3schools.com/tags/tag_meta.asp . [Accessed 12/10/2015].

WebCodeTools [online]. Available at: http://webcodetools.com/. .[Accessed 10/8/2015].

Weideman , Melius. Use of Ethical SEO Methodologies to Achieve Top Rankings in Top Search Engines [online] . In: Proceedings of the 2007 Computer Science and IT Education Conference . Available at: https://www.researchgate.net/publication/250771331_Use_of_Ethical_SEO_Methodologies_to_Achieve_Top_Rankings_in_Top_Search_Engines . [Accessed 22/6/2016].

Wiedeman , Melius and Kritzinger, Wouter. Search Engine Information retrieval [online]: Empirical Research on the Usage of Meta Tags to Enhance Web Site Visibility and Ranking of E-Commerce Web sites.2003. Available at: http://www.web-visibility.co.za/0044-conference-paper-2003-weideman-kritzingersearch-engine-metatags-visibility-e-commerce.pdf .[Accessed 30/5/2016].

Yandex webmaster. Structured data validator [online]. Available at: https://webmaster.yandex.com/microtest.xml . .[Accessed 7/4/2016].

Yu, Liyang. A Developer's Guide to the Semantic Web. 2nd ed. Berlin: Springer, 2014.p. Available at: http://link.springer.com/content/pdf/10.1007%2F978-3-662-43796-4.pdf . [Accessed 24/3/2016].

Zhang ,J. A. . A Study of Metadata Element Co-occurrence. Information Review. Vol. 30, No. 4 (2006). Pp. 428 – 453.

Zhang, Jin and Dimitroff, Alexandra. Internet Search Engines' Response to Metadata Dublin Core Implementation. Journal of Information Science. Vol. 30, No.4 (2004). Pp. 310–320.

Zhang, Jin and Dimitroff, Alexandra. The Impact of Webpage Content Characteristics on Webpage Visibility in Search Engine Results (Part I). Information Processing and Management. Vol. 41 (2005). Pp. 665 - 690. Available at: https://www.semanticscholar.org/paper/The-impact-of-webpage-content-characteristics-on-Zhang-Dimitroff/eb943e7a6307a85d851c3cc23026dcab1151ccac/pdf. [Accessed 28/9/2015].

قائمة الاستهلاليات

1. بؤرة دبلن	DC: Dublin Core
2. مجموعة عناصر بـؤرة دبلـن لمـا وراء البيانات	DCMES: Dublin Core Metadata Element Set
3. مبادرة بؤرة دبلن لما وراء البيانات	DCMI: Dublin Core Metadata Initiative
4. صديق الصديق	FOAF: Friend Of A Friend
5. لغة ترميز النص الفائق	HTML: HyperText Markup Language
6. بروتوكول الرسم البياني	OG: Open Graph Protocol
7. نظام الدوريات المفتوحة	OJS: Open Journal System
8. إطار وصف المصدر	RDF: Resource Description Framework
9. إطار وصف المصدر في الخصائص	RDFa: Resource Description Framework in Attributes
10. تحسين أداء محركات البحث	SEO: Search Engine Optimization
11. صفحة نتائج محرك البحث	SERP:Search Engine Results page
12. المحدد الموحد للمصدر	URI: Uniform Resource Identifier
13. الموقع الموحد للمصدر	URL: Uniform Resource Locator
14. اتحاد الشبكة العنكبوتية العالمية	W3C: World Wide Web Consortium

15. الشبكة العنكبوتية العالمية	WWW: World Wide Web
16. لغة ترميز النص الفائق الموسعة	XHTML: eXtensible HyperText Markup Language
17. لغة الترميز الموسعة	XML: eXtensible Markup Language

مصطلحات الدراسة

1 ما وراء البيانات Metadata

تعني حرفيا بيانات عن البيانات، ويمكن تعريفها بأنها بيانات مهيكلة تصف مصادر المعلومات أو كيانات Objects بهدف تحقيق مجموعة متنوعة من الأغراض[1] تشمل أغراض تحديد الهوية، والاكتشاف والاسترجاع، والتقييم، والإدارة وضبط الإتاحة، والإفادة والحفظ. وعلى الرغم من أن الفهرسة الوصفية سواء كانت في صيغة قواعد الفهرسة الأنجلو أمريكية – وصف المصادر وإتاحتها RDA حاليا – أو في صيغ مارك 21 MARC 21 فإنها تعد نوعا من ما وراء البيانات، إلا أن هذا المصطلح غالبا ما يرتبط في مجال المكتبات والمعلومات بالمصادر الإلكترونية المتاحة في بيئة الشبكات، والخطط غير التقليدية كبؤرة دبلن Dublin Core. وتضع الدراسة الحالية تعريفا لما وراء البيانات ينص على أنها: "بيانات مهيكلة عن أي كيان Object (سواء كان مصدر معلومات – رقمي أو

(1) Reitz, Joan M. ODLIS [online]: Online Dictionary for Library and Information Science. Available at: http://www.abc-clio.com/ODLIS/odlis_m.aspx. [Accessed 29/1/2016].

غير رقمي -، شخص، هيئة، خدمة...إلخ)، وذلك لتحقيق أهداف متنوعة تشمل التعريف والاسترجاع والاستخدام والإدارة.

أما التعريف الإجرائي لما وراء البيانات الذي تتخذه هذه الدراسة فهو: بيانات مهيكلة تصف صفحات العنكبوتية Web Pages المكتوبة بلغة ترميز النص الفائق HTML بهدف تحديد الهوية والاسترجاع والإدارة.

2 محركات البحث Search Engines

هي مواقع تستخدم برامج آلية لتكشيف الصفحات والمواقع المتاحة على الإنترنت في قاعدة بيانات ضخمة، وعندما يقوم أحد الأفراد بالبحث في هذه المحركات، فإنه في واقع الأمر يبحث في هذه القاعدة، ثم تحدد الخوارزمية الخاصة بالمحرك المواقع التي ستظهر في قائمة نتائج البحث، وعادة ما تحاط هذه الخوارزمية أو منطق عمل محرك البحث بالسرية[1].

3 الدوريات الإلكترونية Electronic Journals

هناك اختلاف حول مفهوم الدورية الإلكترونية؛ فهناك من يرى أنها الدورية التي تنشأ في بيئة الإنترنت، وليس لها سابق عهد بالبيئة الورقية، حيث تعتمد على التقنيات الإلكترونية في إنتاجها و الإفادة منها، وعلى شبكات الاتصالات بعيدة المدى في نشرها وتوزيعها، وفي مقابل ذلك هناك من يوسع المجال الدلالي للمصطلح بحيث يشمل الدوريات التي تنشأ في بيئة الإنترنت، والدوريات التي تصدر بالشكلين الورقي والإلكتروني في الوقت نفسه، وتلك الدوريات التي نشأت في بيئة الحبر و الورق ويتم تحويلها إلى الشكل الإلكتروني على نحو راجع.[2].

---

(1) Kok, Ruud. The Value of Search Engines Ranking [online]: The Effect of Search Engine Ranking on a Web Site's Added Value. (Master Thesis). university of Twente, school of Management and Governance, business administration .Available at: http://essay.utwente.nl/59288/1/scriptie_R_Kok.pdf. [Accessed 17/3/2013].

(2) Okerson , Ann. The Electronic Journal: what, whence and when ?. The Public Access Computer Systems Review. Vol. 2, No.1 (1991). Pp. 5-24. & Singleton, A. .The Electronic Journal and Its relatives. Scholarly Publishing , Vol. 13, No. 1 (1981). Pp. 3- 18.
كما استشهد به: حشمت قاسم . الدوريات الإلكترونية التخصصية: تطورها و تحدياتها الاجتماعية والاقتصادية. ص .335. في: حشمت قاسم . الاتصال العلمي في البيئة الإلكترونية . القاهرة: دار غريب للطباعة و النشر و التوزيع، 2005.

وتميل الدراسة إلى توسيع مفهوم الدوريات الإلكترونية ليشمل التعريف الثاني، والذي يتفق مع تعريف قاموس المكتبات وعلم المعلومات على الخط المباشر ODLIS [1] الذي ينص على أن الدورية الإلكترونية هي مقابل رقمي لدورية مطبوعة أو دورية رقمية دون نظير مطبوع، تتاح من خلال العنكبوتية العالمية أو من خلال البريد الإلكتروني أو من خلال أي وسيلة أخرى من وسائل الإتاحة على الإنترنت.

4 العنكبوتية العالمية World Wide Web

شبكة عالمية من خوادم الإنترنت Internet Servers توفر الإتاحة إلى صفحات ومواد مكتوبة بلغة ترميز النص الفائق (Hypertext Markup Language (HTML بما يجعل محتوى هذه الصفحات مترابط. وقام بتصميم العنكبوتية سير "تيم بيرنرز لي" Sir Tim Berners-Lee عام 1989 [2].

5 العنكبوتية الدلالية Semantic Web

العنكبوتية الدلالية هي امتداد للعنكبوتية الحالية وتكون المعلومات فيها محددة المعنى بما يُمكن الحاسبات والبشر من العمل معا.[3] وبعبارة أخرى جعل المعلومات المتاحة على الشبكة مقروءة ومفهومة للآلات بما يجعلها قادرة على القيام بوظائف متقدمة.

(1) Retiz, Joan M. Op. Cit. Available at: http://www.abc-clio.com/ODLIS/odlis_e.aspx#electronicjournal [Accessed 29/1/2016].

(2) Retiz, Joan M. Op. Cit. Available at http://www.abc-clio.com/ODLIS/odlis_w.aspx#www . [Accessed 29/1/2016].

(3) Berners-Lee, T.; Hendler, J. & Lassila, O.: The Semantic Web .Scientific American, Vol. 284, No. 5 (May 2001). Pp..34-43. As cited in: Brase, Jan. O. Usage of metadata [online]. Elektrotechnik und Informatik der Universt at Hannover, 2005.p28. Available at: https://pdfs.semanticscholar.org/ea90/19ad62ce1ab487881610fe393e9d0c4da6ce.pdf. [Accessed 29/11/2013].

6 إطار وصف المصدر Resource Description Framework (RDF)

يعد هذا الإطار أحد النماذج المفاهيمية التي تقدم تصورا نظريا عن كيفية تحديد ووصف المصادر والعلاقات فيما بينها. وهو أحد الركائز الأساسية للعنكبوتية الدلالية Semantic Web، وتم إنشاءه بواسطة اتحاد العنكبوتية W3C عام 1999 كمعيار لتمثيل ما وراء البيانات التي تصف المصادر أو الكيانات بما يجعلها مفهومة آليا [1]. ويعتمد هذا الإطار في وصف المصادر على جمل Statements وتعرف أيضا بالثلاثيات Triples.

7 بؤرة دبلن Dublin core

أحد أشهر خطط ما وراء البيانات، وتم تصميمها لتسهيل عمليات وصف واكتشاف المصادر الإلكترونية للمعلومات في بيئة الإنترنت[2]. وتتكون هذه الخطة في مستواها البسيط من 15 عنصرا تعرف باسم مجموعة عناصر بؤرة دبلن لما وراء البيانات Dublin Core Metadata Element Set وتشمل:

العنوان Title، المنشئ Creator، الموضوع Subject، الوصف Description، الناشر Publisher، المساهم Contributor، التاريخ Date، النوع Type، الشكل Format، المحدد Identifier، المصدر Source، اللغة Language، العلاقة Relation، التغطية Coverage، الحقوق Rights [3]. ومع ظهور إطار وصف المصدر RDF أصبحت بؤرة دبلن من أكثر المفردات Vocabularies استخداما مع هذا الإطار[4].

---

(1) Yu, Liyang. A Developer's Guide to the Semantic Web [online]. 2nd ed. Berlin: Springer, 2014.p. 24 Available at: http://link.springer.com/content/pdf/10.1007%2F978-3-662-43796-4.pdf. [Accessed 24/3/2016]/

(2) Retiz, Joan M. Op. Cit. Available at: http://www.abc-clio.com/ODLIS/odlis_d.aspx#dublincore .[Accessed 291/2016].

(3) Dublin Core Metadata Element Set, Version 1.1[online]. 14-06-2012. Available at: http://www.dublincore.org/documents/dces/. [Accessed 21/3/2016].

(4) DCMI Metadata Basics [online]. Available at: http://dublincore.org/metadata-basics/. [Accessed 19/10/2013].

8 المفردات Vocabularies

مجموعة من المفردات التي تستخدم لغرض محدد. ويمكن أن تتراوح ما بين المفردات البسيطة مثل مجموعة بـؤرة دبلن لما وراء البيانات Dublin Core Metadata Element Set، أو المفردات المعقدة التي تحتوي على آلاف المفردات مثل المستخدمة في مجال الرعاية الصحية لوصف الأمراض، وأعراضها، وطرق العلاج. وهذا المصطلح يتداخل مع مصطلح الأنطولوجيا Ontology [1].

9 إطار وصف المصدر في الخصائص Resource Description Framework In attributes (RDFa)

أحد المعايير الصادرة عن اتحاد العنكبوتية W3C، الذي يوفر مجموعة من الخصائص attributes في لغة ترميز النص الفائق HTML تسمح بجعل المعلومات المتاحة على العنكبوتية مقروءة آليا[2] عن طريق إدراج جمل إطار وصف المصدر RDF Statements.

10 سكيما Schema.org

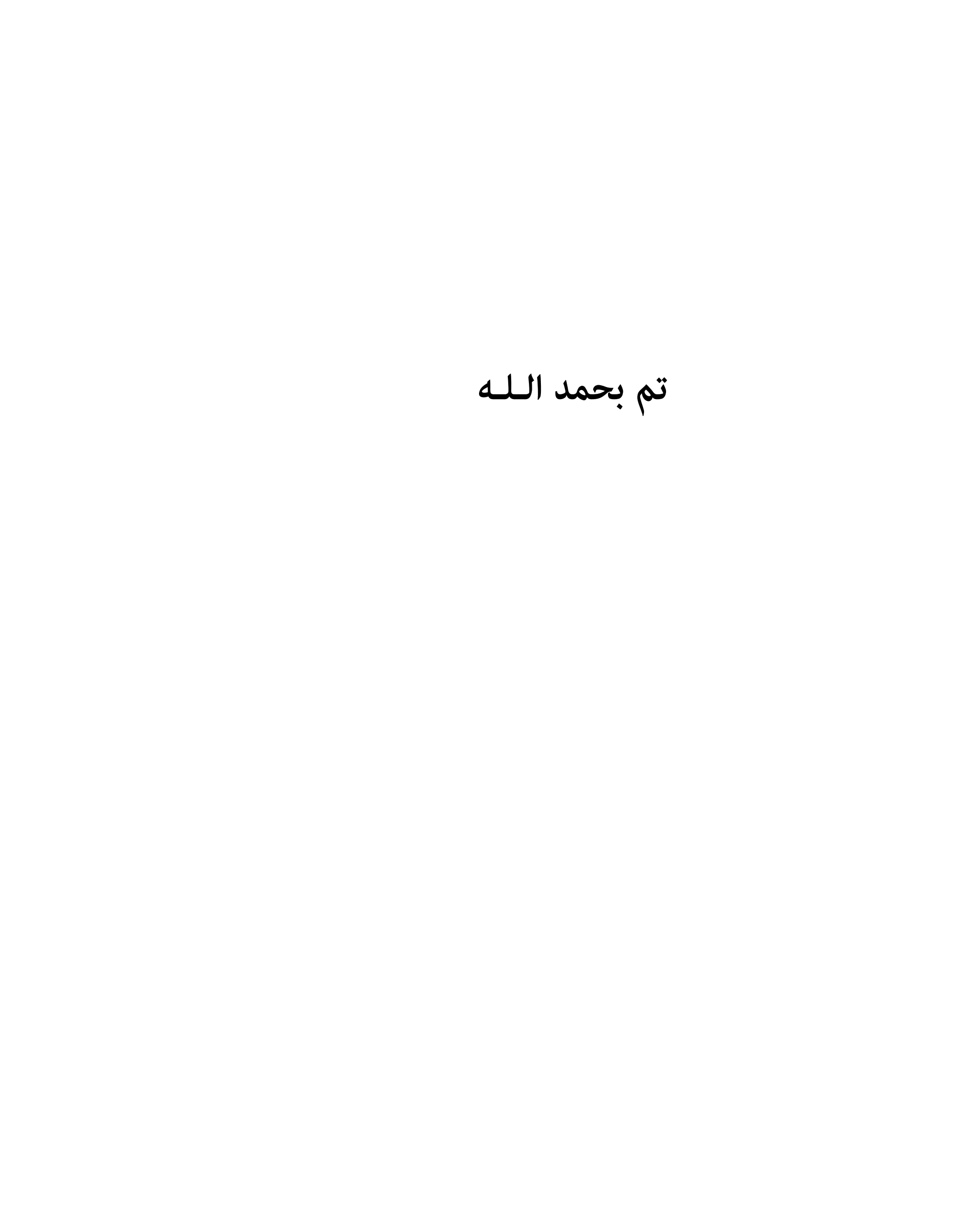
أحد المفردات المستخدمة في بيئة العنكبوتية الدلالية والمستخدمة لجعل البيانات مفهومة آليا. وهي توفر مجموعة من المفردات التي يمكن استخدامها لترميز صفحات العنكبوتية بطريقة تجعلها مفهومة لمحركات البحث الرئيسية مثل جوجل Google، وبينج Bing (مايكروسوفت)، وياهو Yahoo، ومحرك البحث الروسي ياندكس Yandex.[3]

(1) W3C. Linked Data Glossary [online]: W3C Working Group Note 27 June 2013. Available at: https://www.w3.org/TR/2013/NOTE-ld-glossary-20130627/. [Accessed 30/3/2016].

(2) Ibid.

(3) Schema.org. Getting Started - schema.org [online]. Available at: https://schema.org/docs/gs.html . [Accessed 27/5/2016].

تم بحمد الـلـه



تأثير ما وراء البيانات على كفاءة استرجاع مقالات الدوريات الالكترونية العربية



هذا الكتاب

تنظيم المعلومات في بيئة العنكبوتية العاملية واحد من أهم موضوعات علم المكتبات والمعلومات المعاصر .

ونظرا للدور الذي تؤديه ما وراء البيانات « الميتاداتا » تنظيم المعلومات في بيئة العنكبوتية العاملية، فقد كانت وما زالت محوراً للعديد من البحوث والدراسات العربية التي تتناول قضايا تتعلق بخصائصها ووظائفها ولغات التكويد والمعايير المستخدمة في إنشائها، الخ. ومن بين أهم هذه القضايا، القضية التي يتناولها هذا الكتاب وهي قياس تأثير ما وراء البيانات على كفاءة استرجاع مصادر المعلومات العربية من خلال محركات البحث على العنكبوتية العالمية.

ولقد اختارت الباحثة المنهج التجريبي للتحقق من الفروض التي وضعتها لضبط مسار الدراسة والوصول إلى نتائج موضوعية ذات دلالة كمية محسوسة قابلة للمقارنة، بالاضافة إلى معالجة محكمة تتميز بالدقة والوضوح والإيجاز، واستخدام واع للإحصاء الوصفي والاستدلالي.

وهذا الكتاب يسهم في تنمية الوعي بمدى تأثير ما وراء البيانات المعيارية على تحسين كفاءة استراجاع مصادر المعلومات العربية، والمساعدة فى تطوير استخدامها واستثمارها في قواعد البيانات والمكتبات والمستودعات الرقمية العربية. وهو إضافة ثرية للمكتبة العربية في الموضوع نرجو أن يفيد منه المعنيون بالعنكبوتية العاملية ومحركات البحث بصفة عامة، والمعنيون بمصادر المعلومات العربية على العنكبوتية العالمية ، وكذلك المعنيون بعمليات الرقمنة ونظم استرجاع المعلومات في الوطن العربي بصفة خاصة.

والله ولى التوفيق ،،

الناشر

عبدالحي أحمد فؤاد

ISBN:978-977-358-377-4

9 789773 583774

دار الفجر للنشر والتوزيع

4 شارع هاشم الأشقر - النزهة الجديدة - القاهرة

تليفون: 26246252 فاكس : 26246265

www.daralfajr.com info@daralfajr.com